IC17 Rec'd PCT/PTO ATTORNEY'S DOCKET NUMBER

110048

U.S. APPLICATION NO.

(1390 REV. 5-93) US DEPT. OF COMMERCE PATENT & TRADEMARK OFFICE

Other items or information:

16.

TRANSMITTAL LETTER TO THE UNITED STATES **DESIGNATED/ELECTED OFFICE**

(DO/EO/US) CONCERNING A FILING **UNDER 35 U.S.C. 371**

(if known, sec 37 C.F.R.1.5)

09/869927

INTERNATIONAL APPLICATION NO. PRIORITY DATE CLAIMED INTERNATIONAL FILING DATE PCT/FR00/00144 January 21, 2000 January 21, 1999 TITLE OF INVENTION ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT APPLICANT(S) FOR DO/EO/US Glaucia PARAHOS-BACCALA, Francois MALLET, Cecile VOISSET Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information: This is a FIRST submission of items concerning a filing under 35 U.S.C. 371. 1. 2. This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371. 3. This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1). A proper Demand for International Preliminary Examination was made by the 19th month from the earliest 4. claimed priority date. A copy of the International Application as filed (35 U.S.C. 371(c)(2)) a. is transmitted herewith (required only if not transmitted by the International Bureau). = b. 🛛 has been transmitted by the International Bureau. c. is not required, as the application was filed in the United States Receiving Office (RO/US) ā A translation of the International Application into English (35 U.S.C. 371(c)(2)). 6. 7. Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)) a. are transmitted herewith (required only if not transmitted by the International Bureau). have been transmitted by the International Bureau. have not been made; however, the time limit for making such amendments has NOT expired. C. have not been made and will not be made. 8. A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)). 9. An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)). 10. A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)). Items 11. to 16. below concern other document(s) or information included: An Information Disclosure Statement under 37 CFR 1.97 and 1.98. 12. An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included. 13. \bowtie A FIRST preliminary amendment. A SECOND or SUBSEQUENT preliminary amendment. 14. A substitute specification. 15. Entitlement to small entity status is hereby asserted.

JC18 Rec'd PCT/PTO 0 9 JUL 2001 U.S. APPLICATION NO INTERNATIONAL APPLICATION NO. ATTORNEY'S DOCKET NUMBER PCT/FR00/00144 110048 The following fees are submitted: **CALCULATIONS** PTO USE ONLY Basic National fee (37 CFR 1.492(a)(1)-(5)): Search Report has been prepared by the EPO or JPO \$860.00 International preliminary examination fee paid to USPTO (37 CFR1.482)\$690.00 No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2))\$710.00 Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO......\$1,000.00 International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4)\$ 100.00 **ENTER APPROPRIATE BASIC FEE AMOUNT =** \$860.00 Surcharge of \$130.00 for furnishing the oath or declaration later than 20 [] 30 months from the earliest claimed priority date (37 CFR Number Claims Number Filed Extra Rate **Total Claims** 18 - 20 =0 X \$ 18.00 \$ Independent Claims 5 - 3 = 2 X \$80.00 \$160.00 Multiple dependent claim(s)(if applicable) + \$270.00 \$ TOTAL OF ABOVE CALCULATIONS = \$1020.00 Reduction by 1/2 for filing by small entity, if applicable. \$ SUBTOTAL = \$1020.00 Processing fee of \$130.00 for furnishing the English translation later \$ than 20 30 month from the earliest claimed priority date (37 CFR 1.492(f)). TOTAL NATIONAL FEE = \$1020.00 Amount to be refunded Charged Check No. 120730 in the amount of \$1020.00 to cover the above fees is enclosed. a. Please charge my Deposit Account No. ____ in the amount of \$___ to cover the above fees. A duplicate copy b. of this sheet is enclosed. The Director is hereby authorized to charge any additional fees which may be required, or credit any overpayment, C. to Deposit Account No. 15-0461. A duplicate copy of this sheet is enclosed. NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status. SEND ALL CORRESPONDENCE TO: OLIFF & BERRIDGE, PLC P.O. Box 19928 Alexandria, Virginia 22320 William P. Berridge REGISTRATION NUMBER: 30,024

NAME: Joel S. Armstrong

REGISTRATION NUMBER: 36,430

July 9, 2001

Date:

IJ

D

J.

W M

Q

PTO/PCT Rec'd 22 OCT 2001

02100 25 #Y

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of

Glaucia PARANHOS-BACCALA et al.

BOX: SEQUENCE

Application No.: 09/869,927

Filed: August 17, 2001

Docket No.: 110048

For:

ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN

AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT

SUPPLEMENTAL PRELIMINARY AMENDMENT

Director of the U.S. Patent and Trademark Office Washington, D.C. 20231

Sir:

In response to the Notification of Missing Requirements mailed August 21, 2001, please amend the above-identified patent application as follows:

IN THE SPECIFICATION:

At the end of the application, please replace the current Sequence Listing with the attached paper and computer-readable Sequence Listing.

Page 4, line 13 - page 5, line 18, delete current paragraph and insert therefor:

Location of the clones on the reconstructed genomic RNA sequence:

cl.6A2 (1321 bp) 1-1325;

cl.PH74 (535+2229= 2764 bp) 72-606 and 5353-7582;

cl.24.4 (491+1457= 1948 bp); 115-606 and 5353-6810;

cl.44.4 (2372 bp) 115-2496;

cl.PH7 (369+297= 666 bp) 237-606 and 7017-7313; cl.6A1 (2938 bp) 586-3559;

```
cl.Pi5T (2785+566= 3351 bp) 2747-5557 and 7017-7582;
```

cl.7A16 (1422 bp) 2908-4337;

cl.Pi22 (317+1689= 2006 bp) 3957-4273 and 4476-6168;

cl.C4C5 (1116 bp) 6467-7582

5'LTR 1..120

/note="R of 5'LTR (5' end uncertain)"

121..575

/note="U5 of 5'LTR"

misc.

579..596

/note="PBS, primer binding site, for tRNA-W"

misc.

606

/note="splice junction (splice donor site ATCCAAAGTG-GTGAGTAATA

(SEQ ID NO: 32) and splice acceptor site CTTTTTCAG-ATGGGAAACG

(SEQ ID NO: 33), clone RG083M05, GenBank accession AC000064)"

misc.

5353

/note="splice acceptor site for ORF1 (env)"

misc.

5560

/note="splice donor site"

ORF

5581..7194

/note="ORF1 env 538 AA"

/product-="envelope"

misc.

7017

/note="splice acceptor site for ORF2 and ORF3"

ORF

7039..7194

/note="ORF2 52 AA"

ORF 7112..7255

/note="ORF3 48 AA"

misc.

7244..7254

/note="PPT, polypurine tract"

3'LTR 7256..7582

/note="U3-R of 3' LTR (U3-R junction undetermined)

misc.

7563..7569

polyadenylation signal

IN THE CLAIMS:

Please cancel claims 17 and 18 without prejudice to or disclaimer of the subject matter contained therein.

Please replace claim 12 as follows:

12. (Amended) Method for studying and/or monitoring T-cell proliferation in vitro, according to which the T cells from a patient are brought into contact with synthetic peptides belonging to SEQ ID NO. 31.

Please add new claims 19-31 as follows:

- --19. Method according to claim 10, characterized in that the biological sample is a biological fluid chosen from serum, plasma, synovial fluid and urine.--
- --20. Method for studying and/or monitoring T-cell proliferation in vitro, according to which the T cells from a patient are brought into contact with transcription/translation products as obtained according to the method of claim 19.--
- --21. Reagent for detecting, in a biological sample, an autoimmune disease or monitoring pregnancy, comprising at least one transcription/translation product as obtained according to the method of claim 19.--

- --22. Reagent for detecting, in a biological sample, an autoimmune disease or monitoring pregnancy, comprising at least one synthetic peptide belonging to SEQ ID NO: 31.--
- --23. Reagent for detecting, in a biological sample, an autoimmune disease or monitoring pregnancy, comprising at least one protein according to claim 14.--
- --24. A method for detecting susceptibility to an autoimmune disease or monitoring pregnancy of a patient, comprising bringing a biological sample of said patient into contact with at least one fragment according to claim 1.--
- --25. The method of claim 24, wherein said autoimmune disease is multiple sclerosis.--
- --26. A method for detecting susceptibility to an autoimmune disease or monitoring pregnancy of a patient, comprising bringing a biological sample of said patient into contact with at least one transcription/translation product as obtained according to the method of claim 19.--
- --27. The method of claim 26, wherein said autoimmune disease is multiple sclerosis.--
- --28. A method for detecting susceptibility to an autoimmune disease or monitoring pregnancy of a patient, comprising bringing a biological sample of said patient into contact with at least one synthetic peptide belonging to SEQ ID NO: 31.--
- --29. The method of claim 28, wherein said autoimmune disease is multiple sclerosis.--
- --30. A method for detecting susceptibility to an autoimmune disease or monitoring pregnancy of a patient, comprising bringing a biological sample of said patient into contact with at least one protein according to claim 14.--

--31. The method of claim 30, wherein said autoimmune disease is multiple sclerosis.--

REMARKS

Claims 1-16 and 19-31 are pending. Claims 17 and 18 are canceled; claim 12 is amended; and claims 19-31 are added herein.

Claims 19-23 are directed to subject matter deleted by the previous removal of multiple dependencies from the claims. Claims 24-31 are directed to subject matter corresponding to canceled use claims 17 and 18. Claim 12 is amended to remove an

dependency. However, the subject matter deleted from the claim has been

n of claim 20.

r copy and computer-readable copy of the Sequence Listing are vith 37 C.F.R. §§1.821-1.825. The contents of the paper copy and of the Sequence Listing are the same. No new matter is added.

Support for the information provided in the Sequence Listing can be found in the original

Sequence Listing and in the specification at page 4.

Prompt and favorable consideration on the merits is respectfully requested.

Respectfully submitted,

William P. Berridge ((Registration No. 30,024

Melanie L. Mealy Registration No. 40,085

WPB:MLM/jca

Attachments:

Appendix

Sequence Listing (paper and computer-readable copies)

Date: October 22, 2001

OLIFF & BERRIDGE, PLC P.O. Box 19928 Alexandria, Virginia 22320 Telephone: (703) 836-6400 DEPOSIT ACCOUNT USE
AUTHORIZATION
Please grant any extension
necessary for entry;
Charge any fee due to our
Deposit Account No. 15-0461

APPENDIX

Changes to Specification:

The Sequence Listing is replaced.

The following is a marked-up version of the amended paragraph:

Page 4, line 13 - page 5, line 18:

Location of the clones on the reconstructed genomic RNA sequence:

cl.6A2 (1321 bp) 1-1325;

cl.PH74 (535+2229= 2764 bp) 72-606 and 5353-7582;

cl.24.4 (491+1457= 1948 bp); 115-606 and 5353-6810;

cl.44.4 (2372 bp) 115-2496;

cl.PH7 (369+297= 666 bp) 237-606 and 7017-7313; cl.6A1 (2938 bp) 586-3559;

cl.Pi5T (2785+566= 3351 bp) 2747-5557 and 7017-7582;

cl.7A16 (1422 bp) 2908-4337;

cl.Pi22 (317+1689= 2006 bp) 3957-4273 and 4476-6168;

cl.C4C5 (1116 bp) 6467-7582

5'LTR 1..120

/note="R of 5'LTR (5' end uncertain) [sie]"

121..575

/note="U5 of 5'LTR"

misc. 579..596

/note="PBS, primer binding site, for tRNA-W"

misc. 606

/note="splice junction (splice donor site ATCCAAAGTG-GTGAGTAATA

(SEQ ID NO: 32) and splice acceptor site CTTTTTTCAG-ATGGGAAACG_

(SEQ ID NO: 33), clone RG083M05, GenBank accession AC000064)"

misc. 5353

/note="splice acceptor site for ORF1 (env)"

misc. 5560

/note="splice donor site"

ORF 5581..7194

/note="ORF1 env 538 AA"

/product-="envelope"

misc. 7017

/note="splice acceptor site for ORF2 and ORF3"

ORF 7039..7194

/note="ORF2 52 AA"

ORF 7112..7255

/note="ORF3 48 AA"

misc. 7244..7254

/note="PPT, polypurine tract"

3'LTR 7256..7582

/note="U3-R of 3' LTR (U3-R junction undetermined)

misc. 7563..7569

polyadenylation signal

Changes to Claims:

Claims 17 and 18 are canceled.

Claims 19-31 are added.

The following is a marked-up version of the amended claim:

12. (Amended) Method for studying and/or monitoring T-cell proliferation in vitro, according to which the T cells from a patient are brought into contact with either transcription/translation products, as obtained according to the method of claim 11, or synthetic peptides belonging to SEQ ID NO. 31.

SEQUENCE LISTING

<110> PARANHOS-BACCALA, Glaucia

MALLET, Francois

VOISSET, Cecile

 $<\!120\!>$ ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT

<130> 110048

<140> 09/869,927

<141> 2001-08-17

<150> PCT/FR00/00144

<151> 2000-01-21

<150> FR 99/00888

<151> 1999-01-21

<160> 33

<170> PatentIn version 3.1

<210> 1

<211> 1511

<212> DNA

<213> Homo sapiens

<400> 1

cctagaacgt attctggaga attgggacca atgtgacact cagacgctaa gaaagaaacg 60 atttatattc ttctgcagta ccgcctggcc acaatatcct cttcaaggga gagaaacctg 120 gcttcctgag ggaagtataa attataacat catcttacag ctagacctct tctgtagaaa 180 ggagggcaaa tggagtgaag tgccatatgt gcaaactttc ttttcattaa gagacaactc 240 acaattatgt aaaaagtgtg gtttatgccc tacaggaagc cctcagagtc cacctcccta 300 ccccagcgtc ccctccccga ctccttcctc aactaataag gacccccctt taacccaaac 360 ggtccaaaag gagatagaca aaggggtaaa caatgaacca aagagtgcca atattccccg 420 attatgcccc ctccaagcag tgagaggagg agaattcggc ccagccagag tgcctgtacc 480 tttttctctc tcagacttaa agcaaattaa aatagaccta ggtaaattct cagataaccc 540 tgacggctat attgatgttt tacaagggtt aggacaatcc tttgatctga catggagaga 600 tataatgtta ctactaaatc agacactaac cccaaatgag agaagtgccg ctgtaactgc 660 agcccgagag tttggcgatc tttggtatct cagtcaggcc aacaatagga tgacaacaga 720 ggaaagaaca actcccacag gccagcaggc agttcccagt gtagaccctc attgggacac 780 agaatcagaa catggagatt ggtgccacaa acatttgcta acttgcgtgc tagaaggact 840 gaggaaaact aggaagaagc ctatgaatta ctcaatgatg tccactataa cacagggaaa 900 ggaagaaaat cttactgctt ttctggacag actaagggag gcattgagga agcatacctc 960 cctgtcacct gactctattg aaggccaact aatcttaaag gataagttta tcactcagtc 1020 agctgcagac attagaaaaa acttcaaaag tctgccttag gcccggagca gaacttagaa 1080 accctattta acttggcatc ctcagttttt tataatagag atcaggagga gcaggcgaaa 1140 cgggacaaac gggataaaaa aaaaaggggg ggtccactac tttagtcatg gccctcaggc 1200 aagcagactt tggaggctct gcaaaaggga aaagctgggc aaatcaaatg cctaataggg 1260 ctggcttcca gtgcggtcta caaggacact ttaaaaaaga ttatccaagt agaaataagc 1320 cgcccccttg tccatgcccc ttacgtcaag ggaatcactg gaaggcccac tgccccaggg 1380 gatgaagata ctctgagtca gaagccatta accagatgat ccagcagcag gactgagggt 1440 gcccggggcg agcgccagcc catgccatca ccctcacaga gccccgggta tgtttgacca 1500 ttgagagcca a 1511

<210> 2

<211> 2009

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1940)..(1940)

 $\langle 223 \rangle$ n = a or g or c or t/u

<400> 2 atacgactac tatagggcga attgggccct ctagatgcat gctcgagcgg ccgccagtgt 60 gatggatate tgcagaatte geeetttgte egetgtgete etgateeage gaggegeeea 120 ttgctgctcc caattgggct aaaggcttgc cattgttccc acacggctaa gtgcccgggt 180 tcatcctaat tgagctgaac actagtcact gggttccatg gttctcttcc atgacccacg 240 gcttctaata gagctctaat actcaccaca tggcccaaga ttccattcct tggaatccgt 300 gaggecaaga acceeaggte agagaacaeg aggettgeea eegtettgga agtggeeege 360 cgccatcttg ggagctctgg gagcaaggac ccccagtaa cattttggca accacaaagg 420 gacctccaaa gcgatgggaa acattccccc caaggcaaaa acgcccctaa gatgtattct 480 ggagaattgg gaccaatgtg acactcagac gctaagaaag aaacgattta tattcttctg 540 cagtaccgcc tggccacaat atcctcttca agggagagaa acctggcttc ctgagggaag 600 tataaattat aacatcatct tacagctaga cctcttctgt agaaaggagg gcaaatggag 660 tgaagtgcca tatgtgcaaa ctttcttttc attaagagac aactcacaat tatgtaaaaa 720 gtgtggttta tgccctacag gaagccctca gagtccacct ccctacccca gcgtcccccc 780 cccgactcct tcctcaacta ataaggaccc ccctttaacc caaacggtcc aaaaggagat 840 agacaaaggg gtaaacaatg aaccaaagag tgccaatatt ccccgattat gccccctcca 900 agcagtgaga ggaggagaat tcggcccagc cagagtgcct gtaccttttt ctctctcaga 960 cttaaagcaa attaaaatag acctaggtaa attctcagat aaccctgacg gctatattga 1020 tgttttacaa gggttaggac aatcctttga tctgacatgg agagatataa tgttactact 1080 aaatcagaca ctaaccccaa atgagagaag tgccgctgta actgcagccc gagagtttgg 1140 cgatctttgg tatctcagtc aggtcaacaa taggatgaca acagaggaaa gaacaactcc 1200 cacaggccag caggcagttc ccagtgtaga ccctcattgg gacacagaat cagaacatgg 1260

agattggtgc cacaaacatt tgctaacttg cgtgctagaa ggactgagga aaactaggaa 1320 gaagcctatg aattactcaa tgatgtccac tataacacag ggaaaggaag aaaatcctac 1380 1440 tgcttttctg gacagactaa gggaggcatt gaggaagcat acctccctgt cacctgactc 1500 tattgaaggc caactaatct taaaggataa gtttatcact cagtcagctg cagacattag aaaaaacttc aaaagtctgc cttaggcccg gagcagaact tagaaaccct atttaacttg 1560 gcatcctcag ttttttataa tagagatcag gaggagcagg cgaaacggga caaacgggat 1620 aaaaaaaaa gggggggtcc actactttag tcatggccct caggcaagca gactttggag 1680 gctctggaaa agggaaaagc tgggcaaatc aaatgcctaa tagggctggc ttccagtgcg 1740 gtctacaagg acactttaaa aaagattatc caagtagaaa taagccgccc ccttgtccat 1800 gccccttacg tcaagggaat cactggaagg cccactgccc caggggatga agatactctg 1860 agtcagaagc cattaaccag atgatccagc agcaggactg agggtgcccg gggcgagcgc 1920 cagcccatgc catcaccctn acagagcccc gggtatgctt gaccattgag agccaggagg 1980 ttaactgtct cctggacact ggcgcagcc 2009

<210> 3

<211> 1056

<212> DNA

<213> Homo sapiens

<400> 3 ctagaacgta ttctggagaa ttgggaccaa tgtgacactc agacgctaag aaagaaacga 60 tttatattct tctgcagtac cgcctggcca caatatcctc ttcaagggag agaaacctgg 120 cttcctgagg gaagtataaa ttataacatc atcttacagc tagacctctt ctgtagaaag 180 gagggcaaat ggagtgaagt gccatatgtg caaactttct tttcattaag agacaactca 240 caattatgta aaaagtgtgg tttatgccct acaggaagcc ctcagagtcc acctccctac 300 cccagegtee ecteeegae teetteetea actaataagg acceeettt aacccaaacg 360 gtccaaaagg agatagacaa aggggtaaac aatgaaccaa agagtgccaa tattccccga 420 ttatgccccc tccaagcagt gagaggagga gaattcggcc cagccagagt gcctgtacct 480 ttttctctct cagacttaaa gcaaattaaa atagacctag gtaaattctc agataaccct 540 gacggctata ttgatgtttt acaagggtta ggacaatcct ttgatctgac atggagagat 600

ataato	gttac	tactaaatca	gacactaacc	ccaaatgaga	gaagtgccgc	tgtaactgca	660
gcccga	gagt	ttggcgatct	ttggtatctc	agtcaggcca	acaataggat	gacaacagag	720
gaaaga	acaa	ctcccacagg	ccagcaggca	gttcccagtg	tagaccctca	ttgggacaca	780
gaatca	gaac	atggagattg	gtgccacaaa	catttgctaa	cttgcgtgct	agaaggactg	840
aggaaa	acta	ggaagaagcc	tatgaattac	tcaatgatgt	ccactataac	acagggaaag	900
gaagaa	aatc	ttactgcttt	tctggacaga	ctaagggagg	cattgaggaa	gcatacctcc	960
ctgtca	cctg	actctattga	aggccaacta	atcttaaagg	ataagtttat	cactcagtca	1020
gctgca	gaca	ttagaaaaaa	cttcaaaagt	ctgcct			1056
<210>	4						
<211>	25						
<212>	DNA						
<213>	Homo	sapiens					
<400> cggaca	4 tcca	aagtgatggg	aaacg				25
<210>	5						
<211>	26						
<212>	DNA						
<213>	Homo	sapiens					
<400> ggacago	5 gaaa	gtaagactga	gaaggc				26
<210>	6						
<211>	26						
<212>	DNA						
<213>	Homo	sapiens					
	6						
cctagaa	.cgt a	attctggaga a	attggg				26

<210>	7														
<211>	26														
<212>	DNA														
<213>	Homo	sapi	.ens												
<400>					~~~										26
tggctc	ccaa t	ggto	aaac	a ta	.eeeg										20
<210>	8														
<211>	21														
<212>	DNA														
<213>	Homo	sapi	.ens												
<400>															21
atggga	aacg t	tcc	cccg	ga g											21
<210>	9														
<211>	21														
<212>	DNA														
<213>	Homo	sapi	lens												
<400>	9														21
ggccta	aggc a	agact	itttg	ja a											4.1
<210>	10														
<211>	409														
<212>	PRT														
<213>	Homo	sap:	iens												
<400>	10														
Met Gl	y Ser	Ser		His	His	His	His		Ser	Ser	Gly	Leu		Pro	
1			5					10					15		

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg 20 25 30

Ile Met Gly Asn Ile Pro Pro Lys Ala Lys Thr Pro Leu Arg Cys Ile 35 40 45

Leu Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu Arg Lys Lys Arg 50 55 60

Phe Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr Pro Leu Gln Gly 65 70 75 80

Arg Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr Asn Ile Ile Leu 85 90 95

Gln Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp Ser Glu Val Pro 100 105 110

Tyr Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser Gln Leu Cys Lys 115 120 125

Lys Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser Pro Pro Pro Tyr 130 135 140

Pro Ser Val Pro Pro Pro Thr Pro Ser Ser Thr Asn Lys Asp Pro Pro 145 150 155 160

Leu Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly Val Asn Asn Glu 165 170 175

Pro Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro Leu Gln Ala Val Arg 180 185 190

Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro Phe Ser Leu Ser 195 200 205

Asp Leu Lys Gln Ile Lys Ile Asp Leu Gly Lys Phe Ser Asp Asn Pro 210 215 220

Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln Ser Phe Asp Leu 225 230 235 240

Thr Trp Arg Asp Ile Met Leu Leu Leu Asn Gln Thr Leu Thr Pro Asn 245 250 255

Glu Arg Ser Ala Ala Val Thr Ala Ala Arg Glu Phe Gly Asp Leu Trp 260 265 270

Tyr Leu Ser Gln Val Asn Asn Arg Met Thr Thr Glu Glu Arg Thr Thr 275 280 285

Pro Thr Gly Gln Gln Ala Val Pro Ser Val Asp Pro His Trp Asp Thr 290 295 300

Glu Ser Glu His Gly Asp Trp Cys His Lys His Leu Leu Thr Cys Val 305 310 315 320

Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Pro Met Asn Tyr Ser Met 325 330 335

Met Ser Thr Ile Thr Gln Gly Lys Glu Glu Asn Pro Thr Ala Phe Leu 340 345 350

Asp Arg Leu Arg Glu Ala Leu Arg Lys His Thr Ser Leu Ser Pro Asp 355 360 365

Ser Ile Glu Gly Gln Leu Ile Leu Lys Asp Lys Phe Ile Thr Gln Ser 370 375 380

Ala Ala Asp Ile Arg Lys Asn Phe Lys Ser Leu Pro Lys Leu Ala Ala 385 390 395 400

Ala Leu Glu His His His His His 405

<210> 11

<211> 393

<212> PRT

<213> Homo sapiens

<400> 11

Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Ile Met Gly Asn 1 5 10 15

Ile Pro Pro Lys Ala Lys Thr Pro Leu Arg Cys Ile Leu Glu Arg Ile 20 25 30

Leu Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu Arg Lys Lys Arg 35 40 45

Phe Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr Pro Leu Gln Gly 50 55 60

Arg Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr Asn Ile Ile Leu 65 70 75 80

Gln Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp Ser Glu Val Pro 85 90 95

Tyr Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser Gln Leu Cys Lys 100 105 110

Lys Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser Pro Pro Pro Tyr 115 120 125

Pro Ser Val Pro Ser Pro Thr Pro Ser Ser Thr Asn Lys Asp Pro Pro 130 135 140

Leu Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly Val Asn Asn Glu 145 150 155 160

Pro Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro Leu Gln Ala Val Arg 165 170 175

Gly Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro Phe Ser Leu Ser 180 185 190

Asp Leu Lys Gln Ile Lys Ile Asp Leu Gly Lys Phe Ser Asp Asn Pro 195 200 205

Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln Ser Phe Asp Leu 210 215 220

Thr Trp Arg Asp Ile Met Leu Leu Leu Asn Gln Thr Leu Thr Pro Asn 225 230 235 240

Glu	Arg	Ser	Ala	Ala 245	Val	Thr	Ala		Arg 250	Glu	Phe	Gly	Asp	Leu 255	Trp
Tyr	Leu	Ser	Gln	Ala	Asn	Asn	Ara	Met	Thr	Thr	Glu	Glu	Λrα	Πрν	Πрх

Tyr Leu Ser Gln Ala Asn Asn Arg Met Thr Thr Glu Glu Arg Thr Thr 260 265 270

Pro Thr Gly Gln Gln Ala Val Pro Ser Val Asp Pro His Trp Asp Thr 275 280 285

Glu Ser Glu His Gly Asp Trp Cys His Lys His Leu Leu Thr Cys Val 290 295 300

Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Pro Met Asn Tyr Ser Met 305 310 315 320

Met Ser Thr Ile Thr Gln Gly Lys Glu Glu Asn Leu Thr Ala Phe Leu 325 330 335

Asp Arg Leu Arg Glu Ala Leu Arg Lys His Thr Ser Leu Ser Pro Asp 340 345 350

Ser Ile Glu Gly Gln Leu Ile Leu Lys Asp Lys Phe Ile Thr Gln Ser 355 360 365

Ala Ala Asp Ile Arg Lys Asn Phe Lys Ser Leu Pro Lys Leu Ala Ala 370 375 380

Ala Leu Glu His His His His His His 385 390

<210> 12

<211> 26

<212> DNA

<213> Homo sapiens

<400> 12 ctagaacgta ttctggagaa ttggga

26

<210> 13

<211> 20

<212>	DNA	
<213>	Homo sapiens	
<400> cctaa	13 ggcag acttttgaag	20
<210>	14	
<211>	54	
<212>	DNA	
<213>	Homo sapiens	
<400> tttggt	14 Laata cgactcacta tagggcagcc accatgggaa acgttccccc cgag	54
<210>	15	
<211>	44	
<212>	DNA	
<213>	Homo sapiens	
<400> ttttt	15 tttt tttttttc aggctgcgcc agtgtccagg agac	44
<210>	16	
<211>	20	
<212>	DNA	
<213>	Homo sapiens	
<400> tgtccg	16 ctgt gctcctgatc	20
<210>	17	
<211>	44	
<212>	AND	

<213> Homo sapiens

<400> 17	ttttttttc	aggctgcgcc	agtgtccagg	agac		4.4					
<210> 18											
<211> 678	1										
<212> DNA	1										
<213> Hom	Homo sapiens										
<220>											
<221> mis	c_feature										
<222> (59	4)(594)										
<223> n =	a or g or	c or t/u									
<400> 18											
tcagggatag	ccccatcta	tttggccagg	cattagccca	agacttgagc	cagttctcat	60					
acctggatat	tcttgtcctt	tggtatgcgg	atgatttact	tttagccgcc	cgttcagaaa	120					
ccttgtgcca	tcaagccacc	caagtgctct	taaatttcct	cgccacctgt	ggctacaagg	180					
tttccaaacc	aaaggctcag	ctctgctcac	agcagaaggc	tatttaccct	aaatacttag	240					
ggctgaaatt	atccaaaggc	accagggccc	tcagtgagga	atgtatccag	cctatactgg	300					
cttatcctta	tcccaaaacc	ctaaaacaac	taagaaggtt	ccttggcata	ataggcataa	360					
caggcataac	aggtttctgc	tgaatatgga	ttcccaagta	cggcaaaata	gccagaccat	420					
tatatacact	aattaaggaa	actcagaaag	ccaataccca	tttagtaaga	tggacacctg	480					
aagcagaggc	agctttccag	gccgtaaaga	acaccctaac	ccaagcccca	gtgttaagct	540					
tgccagcggg	gcaagacttt	tctttctgtg	tcacagaaaa	aataggaata	gctntaggag	600					
	ggtccgaggg					660					
atgtagtggc						678					
<210> 19											
<211> 591											

<213>

Homo sapiens

<212> DNA

<213> Homo sapiens

<400> 19 ccatggccat ctacactgaa caagatttat acaatcatgt cgtacctaag ccccacaaca 60 aaagagtacc cattetteet tttgttatea gageaggagt getaggeaga etaggtaetg 120 gcattggcag tatcacaacc tctactcagt tctactacaa actatctcaa gaaataaatg 180 gtgacatgga acaggtcact gactccctgg tcaccttgca agatcaactt aactccctag 240 cagcagtagt ccttcaaaat cgaagagctt tagacttgct aaccgccaaa agaggggaa 300 cctgtttatt tttaggagaa gaacgctgtt attatgttaa tcaatccaga attgtcactg 360 agaaagttaa agaaattcga gatcgaatac aatgtagagc agaggagctt caaaacaccg 420 aacgctgggg cctcctcagc caatggatgc cctgggttct ccccttctta ggacctctag 480 cagctctaat attgttactc ctctttggac cctgtatctt taacctcctt gttaagtttg 540 tctcttccag aattgaagct gtaaagctac agatggtctt acaaatctag a 591 <210> 20 <211> 1321 <212> DNA

<400> 20 caacaatcgg gatataaacc caggcattcg agctggcaac agcagccccc ctttgggtcc

60

cttccctttg tatgggagct gttttcatgc tatttcactc tattaaatct tgcaactgca 120 ctcttctggt ccatgtttct tacggctcga gctgagcttt tgctcaccgt ccaccactgc 180 tgtttgccac caccgcagac ctgccgctga ctcccatccc tctggatcct gcagggtgtc 240 cgctgtgctc ctgatccagc gaagcgccca ttgccgctcc caattgggct aaaggcttgc 300 cattgttcct gcacggctaa gtgcctgggt ttgttctaat tgagctgaac actagtcact 360 gggttccatg gttctcttct gtgacccacg gcttctaata gaactataac acttaccaca 420 tggcccaaga ttccattcct tggaatccgt gaggccaaga actccaggtc agagaatacg 480 aagettgeea eeatettgga ageggeetge taccatettg gaagtggtte accaccatet 540 tgggagctct gtgagcaagg acccccggt aacattttgg caaccacgaa cggacatcca 600

aagtgatggg	aaacgttccc	cgcaagacaa	aaacgcccct	aagacgtatt	ctggaaaatt	660
gggaacaatt	tgaccctcag	acactaagaa	agaaacgact	tatattcttc	tgcagtgccg	720
cctggcactc	ctgagggaag	tataaattat	aacaccatct	tacagctaga	cctcttttgt	780
agaaaaggca	aatggagtga	agtgccataa	gtacaaactt	tcttttcatt	aagagacaac	840
tcacaattat	gtaaaaagtg	tgatttatgc	cctacaggaa	gccttcagag	tctacctccc	900
tatcccagca	tccccgactc	cttccccact	taataaggac	cccccttcaa	cccaaatggt	960
ccaaaaggag	atagacaaaa	gggtaaacag	tgaaccaaag	agtgccaata	ttccccaatt	1020
atgacccctc	caagcagtgg	gaggaagaga	attcggccca	gccagagtgc	atgtgccttt	1080
ttctctccca	gacttaaagc	aaataaaaac	agacttaggt	aaattctcag	ataaccctga	1140
tggctatatt	ggtgttttac	aagggttagg	acaattcttt	gatctgacat	ggagagatat	1200
atatgtcact	gctaaatcag	acactaaccc	caaatgagag	aagtgccacc	ataactgcag	1260
cctgagagtt	tggcgatctc	tggtatctca	gtcaggtcaa	tgataggatg	acaacagagg	1320
a						1321

<210> 21

<211> 2938

<212> DNA

<213> Homo sapiens

<400> 21 caacgacgga catccaaagt gatgggaaac gttccccgca agacaaaaac gcccctaaga 60 cgtattctgg agaattggga ccaatttgac cctcagacac taagaaagaa acgacttata 120 ttcttctgca gtgccgcctg gcactcctga gggaagtata aattataaca ccatcttaca 180 gctagacttc ttttgtagaa aaggcaaatg gagtgaagtg ccataagtac aaactttctt 240 ttcattaaga gacaactcac aattatgtaa aaagtgtgat ttatgcccta caggaagcct 300 tcagagtcta cctccctatc ccagcatccc cgactccttc cccaactaat aaggacccc 360 cttcaaccca aatggtccaa aaggagatag acaaaagggt aaacagtgaa ccaaagagtg 420 ccaatattcc ccaattatga cccctcccaa gcagtgggag gaagagattc ggcccagcca 480 gagtgcatgt gctttttctt ctcccagact taaagcaaat aaaaacagac ttaggtaaat 540 tctcagataa tcctgatggc tatattgatg ttttacaagg gttaggacaa ttctttgatc 600

tgacatggag agatataatg tcactgctaa atcagacact aaccccaaat gagagaagtg 660 ccaccataac tgcagcctga gagtttggcg atctctggta tctcagtcag gtcaatgata 720 ggatgacaac agaggaaaga gatgatcccc acagccagca agcagttccc agtctasacc 780 ctcattgggg acacagaaat cagtaacatg ggagattggt gctgcagaca tttgctaact 840 tgtgtgctac aaggactaag gaaaactacg aagaaaatct acgaattact caatgatgtc 900 caccataaca caggggaagg gaagaaaatc ctactgcctt tctggagaga ctaagggagg 960 cattgaggaa gcgtgcctct ctgtcacctg actcttctga aggccaacta atcttaaagc 1020 gtaagtttat cactcagtca gctgcagaca ttagaaaaaa cttcaaaagt ctgccgtagg 1080 cccggagcaa aacttagaaa ccctattgaa cttggcaacy tcggtttttt ataatagaga 1140 tcaggaggag caggcggaac aggacaaacg ggattaaaaa aaaggccacc gctttagtca 1200 tgaccctcag gcaagtggac tttggaggct ctggaaaagg gaaaagctgg gcaaattgaa 1260 tgcctaatag ggcttgcttc cagtgcggtc tacaaggaca ctttaaaaaa gattgtccaa 1320 gtagaagtaa gccgcccctt cgtccatgcc ccttatttca agggaatcac tggaaggccc 1380 actgccccag gggacaaagg tcttttgagt cagaagccac taaccagatg atccagcagc 1440 aggactgagg gtgcctgggg caagcgccat cccatgccat caccctcaca gagccctggg 1500 tatgcttgac cattgagggc caggaaggtt gtctcctgga cactggtgcg gtcttcttag 1560 tcttactctt ctgtcccgga caactgtcct ccagatctgt cactatctga gggggtccta 1620 agacgggcag tcactagata cttctcccag ccactaagtt atgactgggg agctttattc 1680 ttttcacatg cttttctaat tatgcttgaa agccccacta ccttgttagg gagagacatt 1740 ctagcaaaag caggggccat tatacacctg aacataggag aaggaacacc cgtttgttgt 1800 cccctgcttg aggaaggaat taatcctgaa gtctgggcaa cagaaggaca atatggacga 1860 gcaaagaatg cccgtcctgt tcaagttaaa ctaaaggatt ccacttcctt tccctaccaa 1920 aggcagtacc ccctcagacc caaggcccaa caaggattcc aaaagattgt taaggactta 1980 aaagcccaag gcttagtaaa accatgcata actccctgca gtaattccgt agtggattga 2040 ggaggcacag aaacccagtg gacagtggag ggttagtgca agatctcagg attatcaatg 2100 gaggccgttg tccttttata cccagctgta cctagccctt atactgtgct ttcccaaata 2160 ccagaggaag cagagtggtt tacactcctg gaccttaagg atgccttctt ctgcatccct 2220 gtacatcctg actctcaatt cttgtttgcc tttgaagata cttcaaaccc aacatctcaa 2280 ctcacctgga ctgttttacc ccaagggttc agggatagcc cccatctatt tggccaggca 2340

```
ttagcccaag acttgagcca atcctcatac ctggacactt gtccttcggt aggtggatga
                                                                     2400
tttacttttg gccgcccatt cagaaacctt gtgccatcaa gccacccaag cgctcttcaa
                                                                     2460
tttcctcgct acctgtggct acatggtttc caaaccaaag gctcaactct gctcacagca
                                                                     2520
ggttacttag ggctaaaatt atccaaaggc accagggccc tcagtgagga acacatccag
                                                                     2580
cctatactgg cttatcctca tcccaaaacc ctaaagcaac taaggggatt ccttggcgta
                                                                     2640
ataggtttct gccgaaaatg gattcccagg tttggcgaaa tagccaggtc attaaataca
                                                                     2700
ctaattaagg aaactcagaa agccaatacc catttagtaa gatggacaac tgaagtagaa
                                                                     2760
gtggctttcc aggccctaac ccaagcccca gtgttaagtt tgccaacagg gcaagacttt
                                                                    2820
tcttcatatg tcacagaaaa aacaggaata gctctaggag tccttacaca gatccgaggg
                                                                    2880
atgagettge aacetgtgge gtacetgaet aaggaaattg atgtagtgge aaagggtt
                                                                    2938
```

<210> 22

<211> 1422

<212> DNA

<213> Homo sapiens

<220>

<221> misc feature

<222> (879)..(879)

 $\langle 223 \rangle$ n = a or g or c or t/u

<220>

<221> misc_feature

<222> (1200)..(1200)

<223> n = a or g or c or t/u

<400> 22

tcagggatag ccccatcta tttggccagg cattagccca agacttgagt cagttatcat 60 acctggacac tcttgtcctt cagtatgtgg atgatttact tttagctgcc tgttcagaaa 120 ccttgtgcca tcaagccacc caagcactct taaatttcct cgccacctgt ggctacaagg 180

tttccaaaga gaagctcagc tctgctcaca gcaggttaaa tacttaggac taagattatc 240 caaaggcacc aaggccctca gtgaggaatg tatccagcct atactggctt atcctcatct 300 caaaacccta aagcaactaa gagagttcct tggcataaca ggcttctgcc gaatatggat 360 tccccaggta tggcaaaata gccaggccat tatatacagt aattaaggaa actcagaaag 420 ccaataccca tttaataaga tggatacctg aagccaaagt ggctttccag gcccctaaag 480 aaggccttaa acccaagtcc cagtgttaag cttgccaacg gggcaagact tttctttata 540 catcacagaa aaaaacagaa acagctctgg gagtccttac acaggtccaa gggacgagct 600 tgcaacccat ggcatacctg agtaaggaaa ctgatgtagt ggcaaagggt tggcttcatt 660 gtttatgggt agtggtggca gtagcagttg tagtatctga agcagttaaa ataatacagg 720 ggagagatet tactgtgtgg acateteatg aggtgaacag catacteact getaaaggag 780 acttgtggct gtcagacaac cgtttactta aatatcaggc tctattactt gaaaggccag 840 tgctgcaact gtgcacttgt gcaactctta acccagtcnc atttcttcca gacaatgaag 900 atagaatata actgtcaaca aataatttct caaacctatg ccactcgagg ggaccttcta 960 gaagttccct tgactgatcc tgaccttcaa cttgtatact gatggaagtt cctttgtaga 1020 aaaaggactt caaaagcggg gtatgcagtg gtcagtgata atggaatatt tgaaagtatc 1080 ccctcactcc aggaactagt gcttagctgg cagaactaat agccttcatt ggggcactag 1140 aattaggaga aggaaaaagg gtaaatatat atacagactc tgagtatgct cacctagtcn 1200 tccatgccca tgaggcaata tgcagagaaa gggaattcct aacttccgag ggaacaccta 1260 tcacacatca ggaagccatt aggagattat tactggcagt acagaaacct aaagaggtgg 1320 aagtcttaca ctgctggggt catcagaaag gaaagaaaag ggaaatagaa gggaattgcc 1380 aagcagatat tgaagcaaaa agagctgcaa ggcaggaccc tc 1422

<210> 23

<211> 2006

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (305)..(305)

<223> n = a or g or c or t/u

<400> 23 atgcagtggt cagtgataat ggaatacttg aaagtaatcc cctcactcca ggaactagtg 60 ctcagctagc agaactaata gccctcactt gggcactaga attaggagaa gaaaaaaggg 120 caaatatata tacagactct aaatatgctt acctagtcct ccatgcccat gcagcaatat 180 ggaaagaaag ggaatteeta acttetgaga gaacacetat caaacateag gaagecatta 240 ggaaattatt attggctgta cagaaaccta aagaggtggc agtcttacac tgccggggtc 300 atcanaaagg aaaggaaagg gaaatactt ttgcctgcaa ctatccaatg gaaattactt 360 aaaacccttc atcaaacctt tcacttaggc atcgatagca cccatcaaat ggccaaatca 420 ttatttactg gaccaggcct tttcaaaact atcaagcaaa tattcagggc ctgtgaattg 480 tgccaaaaaa ataatcccct gcctcatcgc caagctcctt caggaaaaca aaaaacaggc 540 cattaccctg aaaaaaactg gcaactgatt ttacccacaa gcccaaacct cagggatttc 600 agtatctact agtctgggta aatactttca cgggttgggc aaaggccttc ccctgtagga 660 cagaaaaggc ccaagaggta ataaaggcac tagttcatga aataattccc agattcggac 720 ttccccgagg cttacagagt gacaatagcc ctgctttcca ggccacagta acccagggag 780 tatcccagge gttaggtata cgatatcact tacactgcgc ctgaaggcca cagtcctcag 840 ggaaggtcga gaaaatgaat gaaatactca aaggacatct aaaaaagcaa acccaggaaa 900 cccacctcac atggcctgct ctgttgccta tagccttaaa aagaatctgc aactttcccc 960 aaaaagcagg acttagccca tacgaaatgc tgtatggaag gcccttcata accaatgacc 1020 ttgtgcttga cccaagacag ccaacttagt tgcagacatc acctccttag ccaaatatca 1080 acaagttctt aaaacattac aaggaaccta tccctgagaa gagggaaaag aactattcca 1140 cccttgtgac atggtattag tcaagtccct tctctctaat tccccatccc tagatacatc 1200 ctgggaagga ccctacccag tcattttatt taccccaact gcggttaaag tggctggagt 1260 ggtcttggat acatcacact tgagtcaaat cctggatact gccaaaggaa cctgaaaatc 1320 caggagacaa cgctagctat tcctgtgaac ctctagagga tttgcgcctg ctcttcaaac 1380 aacaaccagg aggaaagtaa ctaaaatcat aaatccccca tggccctccc ttatcatatt 1440 tttctcttta ctgttctttt accctctttc actctcactg cacccctcc atgccgctgt 1500 atgaccagta gctcccctta ccaagagttt ctatggagaa tgcagcgtcc cggaaatatt 1560

```
gatgccccat cgtataggag tctttctaag ggaaccccca ccttcactgc ccacacccat
                                                                    1620
atgccccgca actgctatca ctctgccact ctttgcatgc atgcaaatac tcattattgg
                                                                    1680
acaggaaaaa tgattaatcc tagttgtcct ggaggacttg gagtcactgt ctgttggact
                                                                    1740
tacttcaccc aaactggtat gtctgatggg ggtggagttc aagatcaggc aagagaaaaa
                                                                    1800
catgtaaaag aagtaatctc ccaactcacc cgggtacatg gcacctctag ccctacaaag
                                                                    1860
gactagatct ctcaaaacta catgaaaccc tccgtaccca tactcgcctg gtaagcctat
                                                                    1920
ttaataccac cctcactggg ctccatgagg tctcggccca aaaccctact aactgttgga
                                                                    1980
tatgcctccc cctgaacttc aagcca
                                                                    2006
```

<210> 24

<211> 1948

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (84)..(84)

 $\langle 223 \rangle$ n = a or g or c or t/u

<220>

<221> misc feature

<222> (193)..(193)

 $\langle 223 \rangle$ n = a or g or c or t/u

<220>

<221> misc_feature

<222> (241)..(241)

<223> n = a or g or c or t/u

<400> 24 actgcactct tctggtccat gtttcttacg gctcgagctg agcttttgct caccgtccac 60 cactgctgtt tgccaccacc gcanacctgc cgctgactcc catccctctg gatcctgcag 120 ggtgtccgct gtgctcctga tccagcgagg cgcccattgc cgctcccaat tgggctaaag 180 gcttgccatt gtncctgcac ggctaagtgc ctgggtttgt tctaattgag ctgaacacta 240 ntcactgggt tccatggttc tcttctgtga cccacggctt ctaatagaac tataacactt 300 accacatggc ccaagattcc attccttgga atccgtgagg gcaagaactc caggtcagag 360 aatacgaggc ttgccaccat cttggaagcg gcctgctacc atcttggaag tggttcacca 420 ccatcttggg agctctgtga gcaaggaccc cccggtaaca ttttggcaac cacgaacgga 480 catccaaagt gatacatcct gggaaggacc ctacccagtc attttatcta ccccaactgc 540 ggttaaagtg gctggagtgg agtcttggat acatcacact tgagtcaaat cctggatact 600 gccaaaggaa cctgaaaatc caggagacaa cgctagctat tcctgtgaac ctctagagga 660 tttgcgcctg ctcttcaaac aacaaccagg aggaaagtaa ctaaaatcat aaatccccat 720 ggccctccct tatcatattt ttctctttac tgttgtttca ccctctttca ctctcactgc 780 accccctcca tgccgctgta tgaccagtag ctccccttac caagagtttc tatggagaat 840 gcagcgtccc ggaaatattg atgccccatc gtataggagt ctttgtaagg gaacccccac 900 cttcactgcc cacacccata tgccccgcaa ctgctatcac tctgccactc tttgcatgca 960 tgcaaatact cattattgga caggaaaaat gattaatcct agttgtcctg gaggacttgg 1020 agtcactgtc tgttggactt acttcaccca aactggtatg tctgatgggg gtggagttca 1080 agatcaggca agagaaaaac atgtaaaaga agtaatctcc caactcaccc gggtacatgg 1140 cacctctagc ccctacaaag gactagatct ctcaaaacta catgaaaccc tccgtaccca 1200 tactcgcctg gtaagcctat ttaataccac cctcactggg ctccatgagg tctcggccca 1260 aaaccctact aactgttgga tatgcctccc cctgaacttc aggccatatg tttcaatccc 1320 tgtacctgaa caatggaaca acttcagcac agaaataaac accacttccg ttttagtagg 1380 acctcttgtt tccaatctgg aaataaccca tacctcaaac ctcacctgtg taaaatttag 1440 caatactaca tacacaacca actcccaatg catcaggtgg gtaactcctc ccacacaaat 1500 agtctgccta ccctcaggaa tattttttgt ctgtggtacc tcagcctatc gttgtttgaa 1560 tggctcttca gaatctatgt gcttcctctc attcttagtg ccccctatgg ccatctacac 1620 tgaacaagat ttatacagtt atgtcatatc taagccccgc aacaaaagag tacccattct 1680 tccttttgtt ataggagcag gagtgctagg tgcactaggt actggcattg gcggtatcac 1740

aacctctact cagttctact acaaactatc tcaagaacta aatggggaca tggaacgggt 1800 cgccgactcc ctggtcacct tgcaagatca acttaactcc ctagcagcag tagtccttca 1860 aaatcgaaga gctttagact tgctaaccgc tgaaagaggg ggaacctgtt tattttagg 1920 ggaagaatgc tgttattatg ttaatcaa 1948

<210> 25

<211> 1136

<212> DNA

<213> Homo sapiens

<400> 25 ccatggccat ctacactgaa caagatttat acagttatgt catatctaag ccccgcaaca 60 aaagagtacc cattetteet tttgttatag gageaggagt getaggtgea etaggtaetg 120 gcattggcgg tatcacaacc tctactcagt tctactacaa actatctcaa gaactaaatg 180 gggacatgga acgggtcgcc gactccctgg tcaccttgca agatcaactt aactccctag 240 cagcagtagt ccttcaaaat cgaagagctt tagactcgct aaccgctgaa agagggggaa 300 cctgtttatt tttaggggaa gaatgctgtt attatgttaa tcaatccgga atcgtcactg 360 agaaagttaa agaaattcga gatcgaatac aacgtagagc agaagagctt cgaaacactg 420 gaccctgggg cctcctcagc caatggatgc cctggattct ccccttctta ggacctctag 480 cagctataat attgctactc ctctttggac cctgtatctt taacctcctt gttaactttg 540 tctcttccag aatcgaagct gtaaaactac aaatggagcc caagatgcag tccaagacta 600 agatctaccg cagacccctg gaccggcctg ctagcccacg atctgatgtt aatgacatca 660 aaggcacccc tcctgaggaa atctcagctg cacaacctct actacgcccc aattcagcag 720 gaagcagtta gagcggtcgt cggccaacct ccccaacagc acttaggttt tcctgttgag 780 atgggggact gagagacagg actagctgga tttcctaggc tgactaagaa tccctaagcc 840 tagctgggaa ggtgaccaca tccaccttta aacacggggc ttgcaactta gttcacacct 900 gaccaatcag agagctcact aaaatgctaa ttaggcaaag acaggaggta aagaaatagc 960 caatcatcta ttgcatgaga gcacagcagg agggacaatg atcgggatat aaacccaagt 1020 cttcgagccg gcaacggcaa ccccctttgg gtcccctccc tttgtatggg agctctgttt 1080 tcatgctatt tcactctatt aaatcttgca gctgcgaaaa aaaaaaaaa aaaaaa 1136

<210> 26 <211> 2782 <212> DNA

<213> Homo sapiens

<400> 26 atgggagetg ttttcatgct atttcactct attaaatctt gcaactgcac tcttctggtc 60 catgtttctt acggctcgag ctgagctttt gctcaccgtc caccactgct gtttgccacc 120 accgcagacc tgccgctgac tcccatccct ctggatcctg cagggtgtcc gctgtgctcc 180 tgatccagcg aagcgcccat tgccgctccc aattgggcta aaggcttgcc attgttcctg 240 cacggctaag tgcctgggtt tgttctaatt gagctgaaca ctagtcactg ggttccatgg 300 ttctcttctg tgacccacgg cttctaatag aactataaca cttaccacat ggcccaagat 360 tccattcctt ggaatccgtg aggccaacga actccaggtc agagaatacg aagcttgcca 420 ccatcttgga agcggcctgc taccatcttg gaagtggttc accaccatct tgggagctct 480 gtgagcaagg accccccggt gacattttgg cgaccaccaa cggacatccc aagtgataca 540 teetgggaag gaeeetaeee agteatttta tetaeeeeaa etgeggttaa agtggetgga 600 gtggagtctt ggatacatca cacttgagtc aaatcctgga tactgccaaa ggaacctgaa 660 aatccaggag acaacgctag ctattcctgt gaacctctag aggatttgcg cctgctcttc 720 aaacaacaac caggaggaaa gtaactaaaa tcataaatcc ccatgggcct cccttatcat 780 attittetet gtagtgttet tteaceetgt tteactetea etgeaceece tecatgeege 840 tgtatgacca gtagctcccc tcacccagag tttctatgga gaatgcagcg tcccggaaat 900 attgatgccc catcgtatag gagtetttet aagggaacee ceacetteae tgeecacaee 960 catatgeece geaactgeta teactetgee actetttgea tgeatgeaaa tacteattat 1020 tggacaggaa aaatgattaa teetagttgt eetggaggae ttggagteae tgtetgttgg 1080 acttacttca cccaaactgg tatgtctgat gggggtggag ttcaagatca ggcaagagaa 1140 aaacatgtaa aagaagtaat ctcccaactc accggggtac atggcacctc tagcccctac 1200 aaaggactag atctctcaaa actacatgaa accctccgta cccatactcg cctggtaagc 1260 ctatttaata ccaccctcac tgggctccat gaggtctcgg cccaaaaccc tactaactgt 1320 tggatatgcc tccccctgaa cttcaggcca tatgtttcaa tccctgtacc tgaacaatgg 1380

aacaacttca gcacagaaat aaacaccact tccgttttag taggacctct tgtttccaat 1440 gtggaaataa cccatacctc aaacctcacc tgtgtaaaat ttagcaatac tacatacaca 1500 accaactccc aatgcatcag gtgggtaact cctcccacac aaatagtctg cctaccctca 1560 ggaatatttt ttgtctgtgg tacctcagcc tatcgttgtt tgaatggctc ttcagaatct 1620 atgtgcttcc tctcattctt agtgccccct atgaccatct acactgaaca agatttatac 1680 agttatgtca tatctaagcc ccgcaacaaa agagtaccca ttcttccttt tgttatagga 1740 gcaggagtgc taggtgcact aggtactggc attggcggta tcacaacctc tactcagttc 1800 tactacaaac tatctcaaga actaaatggg gacatggaac gggtcgccga ctccctggtc 1860 accttgcaag atcaacttaa ctccctagca gcagtagtcc ttcgaaatcg aagagcttta 1920 gacttgctaa ccgctgagag agggggaacc tgtttatttt tagggggaaga atgctgttat 1980 tatgttaatc aatccggaat cgtcactgag aaagttgaag aaattccaga tcgaatacaa 2040 cgtatagcag aggagetteg aaacaetgga eeetggggee teetcageeg atggatgeee 2100 tggattctcc ccttcttagg acctctagca gctataatat tgctactcct ctttggaccc 2160 tgtatctttg acctccttgt taactttgtc tcttccagaa tcgaagctgt gaaactacaa 2220 atggagecca agatgeagte caagaetaag atetacegea gaeecetgga eeggeetget 2280 agcccacgat ctgatgttaa tgacatcaaa ggcacccctc ctgaggaaat ctcagctgca 2340 caacctctac tacgccccaa ttcagcagga agcagttaga gcggtggtcg gccaacctcc 2400 ccaacagcac ttaggttttc ctgttgagat gggggactga gagacaggac tagctggatt 2460 tectaggetg actaagaate ettaageeta ggtgggaagg tgaccacate cacetttaaa 2520 cacggggctt gcaacttagc tcacacctga ccaatcagag agctcactaa aatgctaatt 2580 aggcaaagac aggaggtaaa gaaatagcca atcatttatt gcctgagagc acagcaggag 2640 ggacaatgat cgggatataa acccaagttt tcgagccggc aacggcaacc ccctttgggt 2700 cccctccctt tgtatgggag ctctgttttc atgctatttc actctattaa atcttgcaac 2760 tgcaaaaaa aaaaaaaaa aa 2782

<210> 27

<211> 666

<212> DNA

<213> Homo sapiens

<220> <221> misc feature <222> (119)..(119) $\langle 223 \rangle$ n = a or g or c or t/u <400> 27 60 tgtccgctgt gctcctgatc cagcgaggcg cccattgccg ctcccaattg ggctaaaggc ttgccattgt tcctgcacgg ctaagtgcct gggtttgttc taattgagct gaacactant 120 180 cactgggttc catggttctc ttctgtgacc cacggcttct aatataacta taacacttac 240 cacatggccc aagattccat tccttggaat ccgtgaggcc aagaactcca ggtcagagaa 300 tacgaggett gecaccatet tggaagegge etgetaceat ettggaagtg gtteaceace 360 atcttgggag ctctgtgagc aaggaccccc cggtaacatt ttggcaacca cgaacggaca 420 tccaaagtga atcgaagctg taaaactaca aatggagccc aagatgcagt ccaagactaa gatctaccgc agacccctgg accggcctgc tagcccacga tctgatgtta atgacatcaa 480 540 aggcacccct cctgaggaaa tctcagctgc acaacctcta ctacgcccca attcagcagg 600 aagcagttag agcggtcgtc ggccaacctc cccaacagca cttaggtttt cctgttgaga 660 tgggggactg agagacagga ctagctggat ttcctaggct gactaagaat ccctaagcct 666 agctgg <210> 28 <211> 3372 <212> DNA <213> Homo sapiens <400> 28 60 gacttcccaa ataccagagg aagcagagtg gtttacagtc ctggaccttc aggatgcctt 120 cttctgcatc cctgtacatc ctgactctca attcttgttt gcctttgaag atacttcaaa 180 cccagcatct caactcacct ggactatttt accccaaggg ttcagggata gtccccatct 240 atttggccag gcattagccc aagacttgag ccaatcctca tacctggaca cttgtccttc

ggtaggtgga tgatttactt ttggccgccc attcagaaac cttgtgccat caagccaccc

300

360 aagcgctctt caatttcctc gctacctgtg gctacatggt ttccaaacca aaggctcaac tetgeteaca geaggttact tagggetaaa attateeaaa ggeaceaggg eeeteagtga 420 480 ggaacacate cageetatae tggettatee teateecaaa accetaaage aactaagggg 540 atteettgge gtaataggtt tetgeegaaa atggatteee aggtatggeg aaatageeag 600 gtcattaaat acactaatta aggaaactca gaaagccaat acccatttag taagatggac aactgaagta gaagtggctt tccaggccct aacccaagcc ccagtgttaa gtttgccaac 660 720 agggcaagac ttttgttcat atgtcacaga aaaaacagga atagctctag gagtccttac acagatccga gggatgagct tgcaacctgt ggcacacctg actaaggaaa ttgatgtagt 780 ggcaaagggt tgacctcatt gtttacgggt agtggtggca gtagcagtct tagtatctga 840 agcagttaaa ataatacagg gaagagatct tactgtgtgg acatctcatg atgtgaatgg 900 catactcact gctaaaggag acttgtggct gtcagacaac tgtttactta aatgtcaggc 960 1020 tetattactt gaagggeeag tgetgegaet gtgeaettgt geaactetta acceageeae 1080 atttcttcca gacaatgaag aaaagataaa acataactgt caacaagtaa tttctcaaac 1140 ctatgccact cgaggggacc ttttagaggt tcctttgact gatcccgacc tcaacttgta 1200 tactgatgga agttcctttg tagaaaaagg acttcgaaaa gtggggtatg cagtggtcag 1260 tgataatgga atacttgaaa gtaatcccct cactccagga actagtgctc agctagcaga 1320 actaatagcc ctcacttggg cactagaatt aggagaagaa aaaagggcaa atataataca 1380 gactctaaat atgcttacct agtcctccat gcccatgcag caatatggaa agaaagggaa ttcctaactt ctgagagaac acctatcaaa catcaggaag ccattaggaa attattattg 1440 1500 gctgtacaga aacctagaga ggtggcagtc ttacactgcc ggggtcatca caaaggaaag 1560 gaaagggaaa tacaagagaa ctgccaagca tatattgaag ccaaaagagc tgcaaggcag 1620 gaccctccat tagaaatgct tattaaactt cccttagtat agggtaatcc cttccgggaa 1680 accaagcccc agtactcagc aggagaaaca gaatggggaa cctcacgagg cagttttctc 1740 ccctcgggac ggttagccac tgaagaaggg aaaatacttt tgcctgcaac tatccaatgg 1800 aaattactta aaacccttca tcaaaccttt cacttaggca tcgatagcac ccatcagatg gccaaatcat tatttactgg accaggcctt ttcaaaacta tcaagcagat agtcagggcc 1860 tgtgaagtgt gccagagaaa taatcccctg ccttatcgcc aagctccttc aggagaacaa 1920 agaacaggcc attaccetgg agaagactgg caactgattt tacccacaag cccaaacctc 1980 agggatttca gtatctacta gtctgggtag atactttcac gggttgggca gaggccttcc 2040

cctgtaggac agaaaaggcc caagaggtaa taaaggcact agttcatgaa ataattccca 2100 gattcggact tccccgaggc ttacagagtg acaatagccc tgctttccag gccacagtaa 2160 cccagggagt atcccaggcg ttaggtatac gatatcactt acactgcgcc tgaaggccac 2220 agtcctcagg gaaggtcgag aaaatgaatg aaacactcaa aggacatcta aaaaagcaaa 2280 cccaggaaac ccacctcaca tggcctgttc tgttgcctat agccttaaaa agaatctgca 2340 actttcccca aaaagcagga cttagcccat acgaaatgct gtatgqaaqq cccttcataa 2400 ccaatgacct tgtgcttgac ccaagacagc caacttagtt gcagacatca cctccttagc 2460 caaatatcaa caagttotta aaacattaca aggaacctat cootgagaag aggaaaagaa 2520 tattccaccc aagtgacatg gtattagtca agtcccttcc ctctaattcc ccatccctag 2580 atacatcctg ggaaggaccc tacccagtca ttttatctac cccaactgcg gttaaagtgg 2640 ctggagtgga gtcttggata catcacactt gagtcaaatc ctggatactg ccaaaqgaac 2700 ctgaaaatcc aggagacaac gctagctatt cctgtgaacc tctagaggat ttgcgcctgc 2760 tcttcaaaca acaaccagga ggaaaaatcg aagctgtaaa actacaaatg gagcccaaga 2820 tgcagtccaa gactaagatc taccgcagac ccctggaccg gcctgttagc ccacqatctg 2880 atgttaatga catcaaaggc accectectg aggaaatete agetgeacaa cetetactae 2940 gccccaattc agcaggaagc agttagagcg gtcgtcggcc aacctcccca acagcactta 3000 ggttttcctg ttgagatggg ggactgagag acaggactag ctggatttcc taggctgatt 3060 aagaatccct aagcctagct gggaaggtga ccacatccac ctttaaacac ggggcttgca 3120 acttagetea cacetgacea ateagagage teactaaaat getaattagg caaagacagg 3180 aggtaaagaa atagccaatc atttattgcc tgagagcaca gcaggaggga caatgatcgg 3240 gatataaacc caagttttcg agccggcaac ggcaaccccc tttgggtccc ctccctttgt 3300 atgggagete tgttttcatg ctatttcact ctattaaate ttgcaactge aaaaaaaaa 3360 aaaaaaaaa aa 3372

<210> 29

<211> 2372

<212> DNA

<213> Homo sapiens

```
<220>
<221>
       misc feature
<222>
       (1191)..(1191)
<223> n = a or g or c or t/u
<220>
<221> misc_feature
<222>
       (1213)..(1213)
<223> n = a or g or c or t/u
<220>
<221> misc_feature
<222>
       (2089)..(2089)
\langle 223 \rangle n = a or g or c or t/u
<220>
<221> misc feature
<222>
      (2274)..(2274)
\langle 223 \rangle n = a or g or c or t/u
<400> 29
actgcactct tctggtccat gtttcttacg gctcgagctg agcttttgct caccgtccac
                                                                         60
cactgctgtt tgccaccacc gcagacctgc cgctgactcc catccctctg gatcctgcag
                                                                        120
ggtgtccgct gtgctcctga tccagcgagg cgcccattgc cgctcccaat tgggctaaag
                                                                        180
gcttgccatt gttcctgcac ggctaagtgc ctgggtttgt tctaattgag ctgaacacta
                                                                        240
atcactgggt tccatggttc tcttctgtga cccacggctt ctaatagaac tataacactt
                                                                        300
accacatggc ccaagattcc attccttgga atccgtgagg ccaagaactc caggtcagag
                                                                        360
aatacgaggc ttgccaccat cttggaagcg gcctgctacc gtcttggaag tggttcacca
```

ccatcttggg agctctgtga gcaaggaccc cccggtaaca ttttggcaac caacgacgga

catccaaagt gatgggaaac gttccccgca agacaaaaac gcccctaaga cgtattctgg

420

480

540

agaattggga ccaatttgac cctcagacac taagaaagaa acgacttata ttcttctgca 600 gtgccgcctg gcactcctga gggaagtata aattataaca ccatcttaca gctagacctc 660 ttttgtagaa aaggcaaatg gagtgaagtg ccataagtac aaactttctt ttcattaaga 720 gacaactcac aattatgtaa aaagtgtgat ttatgcccta caggaagcct tcagagtcta 780 cctccctatc ccagcatccc cgactccttc cccaactaat aaggaccccc cttcaaccca 840 aatggtccaa aaggagatag acaaaagggt aaacagtgaa ccaaagagtg ccaatattcc 900 ccaattatga cccctccaag cagtgggagg aagagaattc ggcccagcca gagtgcatgt 960 gcctttttct ctcccagact taaagcaaat aaaaacagac ttaggtaaat tctcagataa 1020 ccctgatggc tatattgatg ttttacaagg gttaggacaa ttctttgatc tgacatggag 1080 agatataatg tcactgctaa atcagacact aaccccaaat gagagaagtg ccaccataac 1140 tgcagcctga gggtttggcg tctctggtat ctcagtcagg tcaatggata nggatgacaa 1200 cagaaggaaa ganaatgatt ccccacaggc cagcaggcag ttcccagtct agaccctcat 1260 tgggacacag aatcagaaca tggagattgg tgctgcagac atttgctaac ttgtgtgcta 1320 gaaggactaa ggaaaactag gaagaagtct atgaattact caatgatgtc caccataaca 1380 cagggaaggg aagaaaatcc tactgccttt ctggagagac taagggaggc attgaggaag 1440 cgtgcctctc tgtcacctga ctcttctgaa ggccaactaa tcttaaagcg taagtttatc 1500 actcagtcag ctgcagacat tagaaaaaac ttcaaaaagtc tgccgtaggc ccggagcaaa 1560 acttagaaac cctattgaac ttggcaacct cggtttttta taatagagat caggaggagc 1620 aggcggaaca ggacaaacgg gattaaaaaa aaggccaccg ctttagtcat gaccctcagg 1680 caagtggact ttggaggctc tggaaaaggg aaaagctggg caaattgaat gcctaatagg 1740 gcttgcttcc agtgcggtct acaaggacac tttaaaaaag attgtccaag tagaagtaag 1800 ccgccccttc gtccatgccc cttatttcaa gggaatcact ggaaggccca ctgccccagg 1860 ggacaaaggt cttttgagtc agaagccact aaccagatga tccagcagca ggactgaggg 1920 tgcctggggc aagcgccatc ccatgccatc accctcacag agccctgggt atgcttgacc 1980 attgagggcc aggaaggttg tctcctggac actggtgcgg tcttcttagt cttactcttc 2040 tgtcccggac aactgtcctc cagatctgtc actattctga gggggtccnt aagacgggca 2100 gtcactagat actttttccc agccactaag ttatgaactg gggagcttta ttctttcac 2160 atgcttttct aattatgctt gaaagcccca ctaccttgtt agggagagac attctagcaa 2220 aagcaggggc cattatacac ctgaacatag gagaaggaac acccgtttgt tgtncccctg 2280

<220>

```
cttgaggaag gaattaatcc tgaagtctgg gcaacagaag gacaatatgg acgagccaaa
                                                                     2340
gaatgcccgt cctgttcaag ttaaactaaa gg
                                                                     2372
<210> 30
<211> 7582
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> (198)..(198)
<223> n = a or g or c or t/u
<220>
<221> misc_feature
<222> (307)..(307)
<223> n = a or g or c or t/u
<220>
<221> misc_feature
<222> (355)..(355)
\langle 223 \rangle n = a or g or c or t/u
<220>
<221> misc feature
<222> (1309)..(1309)
<223> n = a or g or c or t/u
```

```
<221> misc_feature
<222> (1331)..(1331)
\langle 223 \rangle n = a or g or c or t/u
<220>
<221> misc_feature
<222> (2213)..(2213)
<223> n = a or g or c or t/u
<220>
<221> misc_feature
<222> (2398)..(2398)
\langle 223 \rangle n = a or g or c or t/u
<220>
<221> misc_feature
<222> (3787)..(3787)
\langle 223 \rangle n = a or g or c or t/u
<220>
<221> misc feature
<222> (4115)..(4115)
<223> n = a or g or c or t/u
<220>
<221> misc_feature
<222> (4261)..(4261)
\langle 223 \rangle n = a or g or c or t/u
```

<400> 30 caacaatcgg gatataaacc caggcattcg agctggcaac agcagccccc ctttgggtcc 60 cttccctttg tatgggagct gttttcatgc tatttcactc tattaaatct tgcaactgca 120 ctcttctggt ccatgtttct tacggctcga gctgagcttt tgctcaccgt ccaccactgc 180 tgtttgccac caccgcanac ctgccgctga ctcccatccc tctggatcct gcagggtgtc 240 cgctgtgctc ctgatccagc gargcgccca ttgccgctcc caattgggct aaaggcttgc 300 cattgtncct gcacggctaa gtgcctgggt ttgttctaat tgagctgaac actantcact 360 gggttccatg gttctcttct gtgacccacg gcttctaata kaactataac acttaccaca 420 tggcccaaga ttccattcct tggaatccgt gaggscaacg aactccaggt cagagaatac 480 gargettgee accatettgg aageggeetg etacertett ggaagtggtt caccaceate 540 ttgggagctc tgtgagcaag gacccccgg tracattttg gcraccamsr acggacatcc 600 660 maagtgatgg gaaacgttcc ccgcaagaca aaaacgcccc taagacgtat tctggaraat tgggamcaat ttgaccctca gacactaaga aagaaacgac ttatattctt ctgcagtgcc 720 gcctggcact cctgagggaa gtataaatta taacaccatc ttacagctag acytcttttg 780 tagaaaaggc aaatggagtg aagtgccata agtacaaact ttcttttcat taagagacaa 840 ctcacaatta tgtaaaaagt gtgatttatg ccctacagga agccttcaga gtctacctcc 900 ctatcccagc atccccgact ccttccccam ytaataagga cccccttca acccaaatgg 960 tccaaaagga gatagacaaa agggtaaaca gtgaaccaaa gagtgccaat attccccaat 1020 tatgacccct cccaagcagt gggaggaaga gaattcggcc cagccagagt gcatgtgcyt 1080 tttyytctcc cagacttaaa gcaaataaaa acagacttag gtaaattctc agataaycct 1140 gatggctata ttgrtgtttt acaagggtta ggacaattct ttgatctgac atggagagat 1200 atatatgtca ctgctaaatc agacactaac cccaaatgag agaagtgcca ccataactgc 1260 agcctgagrg tttggcgatc tctggtatct cagtcaggtc aatggatang gatgacaaca 1320 gaaggaaaga naatgattcc ccacaggcca gcargcagtt cccagtctas accctcattg 1380 gggacacaga aatcagtaac atgggagatt ggtgctgcag acatttgcta acttgtgtgc 1440 tasaaggact aaggaaaact asgaagaaar tctaygaatt actcaatgat gtccaccata 1500 acacagggga agggaagaaa atcctactgc ctttctggag agactaaggg aggcattgag 1560 gaagcgtgcc tctctgtcac ctgactcttc tgaaggccaa ctaatcttaa agcgtaagtt 1620 tatcactcag tcagctgcag acattagaaa aaacttcaaa agtctgccgt aggcccggag 1680 caaaacttag aaaccctatt gaacttggca acytcggttt tttataatag agatcaggag 1740

gagcaggcgg aacaggacaa acgggattaa aaaaaaggcc accgctttag tcatgaccct 1800 1860 caggcaagtg gactttggag gctctggaaa agggaaaagc tgggcaaatt gaatgcctaa tagggettge ttecagtgeg gtetacaagg acaetttaaa aaagattgte caagtagaag 1920 1980 taagccgccc cttcgtccat gccccttatt tcaagggaat cactggaagg cccactgccc caggggacaa aggtcttttg agtcagaagc cactaaccag atgatccagc agcaggactg 2040 2100 agggtgcctg gggcaagcgc catcccatgc catcaccctc acagagccct gggtatgctt gaccattgag ggccaggaag gttgtctcct ggacactggt gcggtcttct tagtcttact 2160 cttctgtccc ggacaactgt cctccagatc tgtcactatt ctgagggggt ccntaagacg 2220 ggcagtcact agatacttty tcccagccac taagttatga actggggagc tttattcttt 2280 tcacatgctt ttctaattat gcttgaaagc cccactacct tgttagggag agacattcta 2340 gcaaaagcag gggccattat acacctgaac ataggagaag gaacacccgt ttgttgtncc 2400 cctgcttgag gaaggaatta atcctgaagt ctgggcaaca gaaggacaat atggacgagc 2460 caaagaatgc ccgtcctgtt caagttaaac taaaggattc cacttccttt ccctaccaaa 2520 ggcagtaccc cctcagaccc aaggcccaac aaggattcca aaagattgtt aaggacttaa 2580 aagcccaagg cttagtaaaa ccatgcataa ctccctgcag taattccgta gtggattgag 2640 gaggcacaga aacccagtgg acagtggagg gttagtgcaa gatctcagga ttatcaatgg 2700 aggccgttgt ccttttatac ccagctgtac ctagccctta tactgtgmyt tcccaaatac 2760 cagaggaagc agagtggttt acastcctgg accttmagga tgccttcttc tgcatccctg 2820 tacateetga eteteaatte ttgtttgeet ttgaagatae tteaaaceea reateteaae 2880 tcacctggac trttttaccc caagggttca gggatagycc ccatctattt ggccaggcat 2940 tagcccaaga cttgagycar tymtcatacc tggacactct tgtccttcrg takgtggatg 3000 atttactttt rgcygccyrt tcagaaacct tgtgccatca agccacccaa gcrctcttma 3060 atttcctcgc yacctgtggc tacawggttt ccaaacsara rgctcarctc tgctcacagc 3120 aggttaaata cttaggrcta arattatcca aaggcaccar ggccctcagt gaggaayrya 3180 tccagcctat actggcttat cctcatcyca aaaccctaaa gcaactaagr grrttccttg 3240 gcrtaayagg yttctgccga awatggattc cccaggtwtg gcraaatagc caggycatta 3300 watacastaa ttaaggaaac tcagaaagcc aatacccatt tartaagatg gayamctgaa 3360 gymraagtgg ctttccaggc ccctaaagaa ggccttaaac ccaagyccca gtgttaagyt 3420 tgccaacrgg gcaagacttt tsttyatayr tcacagaaaa aaacagraay agctctrgga 3480 gtccttacac agrtccragg gaygagcttg caaccyrtgg cryacctgas taaggaaayt 3540 gatgtagtgg caaagggttg rcytcattgt ttaygggtag tggtggcagt agcagtykta 3600 gtatctgaag cagttaaaat aatacagggr agagatctta ctgtgtggac atctcatgak 3660 gtgaayrgca tactcactgc taaaggagac ttgtggctgt cagacaacyg tttacttaaa 3720 trtcaggctc tattacttga arggccagtg ctgcractgt gcacttgtgc aactcttaac 3780 ccagycncat ttcttccaga caatgaagaa aagataraay ataactgtca acaartaatt 3840 teteaaacet atgeeacteg aggggaeett ytagargtte cyttgaetga teeygaeett 3900 caacttgtat actgatggaa gttcctttgt agaaaaagga cttcgaaaaag yggggtatgc 3960 agtggtcagt gataatggaa tayttgaaag taatcccctc actccaggaa ctagtgctya 4020 gctrgcagaa ctaatagccy tcaytkgggc actagaatta ggagaagraa aaagggyaaa 4080 tatatataca gactetrart atgetyacet agtentecat geceatgmrg caatatgsar 4140 agaaagggaa ttcctaactt cygagrgaac acctatcama catcaggaag ccattaggar 4200 attattaytg gcwgtacaga aacctaraga ggtggmagtc ttacactgcy ggggtcatca 4260 naaaggaaag raaagggaaa tasaagrgaa ytgccaagca katattgaag cmaaaagagc 4320 tgcaaggcag gaccctccat tagaaatgct tattaaactt cccttagtat agggtaatcc 4380 cttccgggaa accaagcccc agtactcagc aggagaaaca gaatggggaa cctcacgagg 4440 cagttttctc ccctcgggac ggttagccac tgaagaaggg aaaatacttt tgcctgcaac 4500 tatccaatgg aaattactta aaacccttca tcaaaccttt cacttaggca tcgatagcac 4560 ccatcaratg gccaaatcat tatttactgg accaggcctt ttcaaaacta tcaagcarat 4620 4680 aktcagggcc tgtgaaktgt gccararaaa taatcccctg cctyatcgcc aagctccttc aggaraacaa araacaggcc attaccctgr araaractgg caactgattt tacccacaag 4740 cccaaacctc agggatttca gtatctacta gtctgggtar atactttcac gggttgggca 4800 raggccttcc cctgtaggac agaaaaggcc caagaggtaa taaaggcact agttcatgaa 4860 ataattccca gattcggact tccccgaggc ttacagagtg acaatagccc tgctttccag 4920 gccacagtaa cccagggagt atcccaggcg ttaggtatac gatatcactt acactgcgcc 4980 tgaaggccac agtcctcagg gaaggtcgag aaaatgaatg aaayactcaa aggacatcta 5040 aaaaagcaaa cccaggaaac ccacctcaca tggcctgytc tgttgcctat agccttaaaa 5100 agaatetgea aettteeeca aaaageagga ettageeeat aegaaatget gtatggaagg 5160 cccttcataa ccaatgacct tgtgcttgac ccaagacagc caacttagtt gcagacatca 5220 cctccttagc caaatatcaa caagttctta aaacattaca aggaacctat ccctgagaag 5280 agggaaaaga actattccac ccwwgtgaca tggtattagt caagtccctt cyctctaatt 5340 ccccatccct agatacatcc tgggaaggac cctacccagt cattttatyt accccaactg 5400 cggttaaagt ggctggagtg gagtcttgga tacatcacac ttgagtcaaa tcctggatac 5460 tgccaaagga acctgaaaat ccaggagaca acgctagcta ttcctgtgaa cctctagagg 5520 atttgcgcct gctcttcaaa caacaaccag gaggaaagta actaaaatca taaatccccc 5580 atggscctcc cttatcatat ttttctctkt astgttsttt yaccctsttt cactctcact 5640 gcacccctc catgccgctg tatgaccagt agctccccty accmagagtt tctatggaga 5700 atgcagcgtc ccggaaatat tgatgcccca tcgtatagga gtctttstaa gggaaccccc 5760 accttcactg cccacaccca tatgccccgc aactgctatc actctgccac tctttgcatg 5820 catgcaaata ctcattattg gacaggaaaa atgattaatc ctagttgtcc tggaggactt 5880 ggagtcactg tctgttggac ttacttcacc caaactggta tgtctgatgg gggtggagtt 5940 caagatcagg caagagaaaa acatgtaaaa gaagtaatct cccaactcac csgggtacat 6000 ggcacctcta gcccctacaa aggactagat ctctcaaaac tacatgaaac cctccgtacc 6060 catactcgcc tggtaagcct atttaatacc accctcactg ggctccatga ggtctcggcc 6120 caaaacccta ctaactgttg gatatgcctc cccctgaact tcargccata tgtttcaatc 6180 cctgtacctg aacaatggaa caacttcagc acagaaataa acaccacttc cgttttagta 6240 ggacctcttg tttccaatst ggaaataacc catacctcaa acctcacctg tgtaaaattt 6300 agcaatacta catacacaac caactcccaa tgcatcaggt gggtaactcc tcccacaca 6360 atagtetgee tacceteagg aatatttttt gtetgtggta ceteageeta tegttgtttg 6420 aatggctctt cagaatctat gtgcttcctc tcattcttag tgcccccyat grccatctac 6480 actgaacaag atttatacag ttatgtcata tctaagcccc gcaacaaaag agtacccatt 6540 cttccttttg ttataggagc aggagtgcta ggtgcactag gtactggcat tggcggtatc 6600 acaaceteta etcagiteta etacaaaeta teteaagaae taaatgggga eatggaaegg 6660 gtcgccgact ccctggtcac cttgcaagat caacttaact ccctagcagc agtagtcctt 6720 craaatcgaa gagctttaga ctygctaacc gctgaragag ggggaacctg tttattttta 6780 ggggaagaat gctgttatta tgttaatcaa tccggaatcg tcactgagaa agttraagaa 6840 attcsagatc gaatacaacg takagcagar gagcttcgaa acactggacc ctggggcctc 6900 ctcagccrat ggatgccctg gattctcccc ttcttaggac ctctagcagc tataatattg 6960

ctactcctct ttggaccctg tatctttrac ctccttgtta actttgtctc ttccagaatc 7020 gaagctgtra aactacaaat ggagcccaag atgcagtcca agactaagat ctaccgcaga 7080 cccctggacc ggcctgytag cccacgatct gatgttaatg acatcaaaqg cacccctcct 7140 gaggaaatct cagctgcaca acctctacta cgccccaatt cagcaggaag cagttagagc 7200 ggtsgtcggc caacctcccc aacagcactt aggttttcct gttgagatgg gggactgaga 7260 gacaggacta gctggatttc ctaggctgay taaqaatccy taaqcctaqs tqqqaaqqtq 7320 accacatcca cctttaaaca cggggcttgc aacttagytc acacctgacc aatcagagag 7380 ctcactaaaa tgctaattag gcaaagacag gaggtaaaga aatagccaat catytattgc 7440 mtgagagcac agcaggaggg acaatgatcg ggatataaac ccaagtyttc gagccggcaa 7500 cggcaacccc ctttgggtcc cctccctttg tatgggaget ctgttttcat gctatttcac 7560 tctattaaat cttgcarctg cr 7582

<210> 31

<211> 363

<212> PRT

<213> Homo sapiens

<400> 31

Met Gly Asn Ile Pro Pro Lys Ala Lys Thr Pro Leu Arg Cys Ile Leu 1 5 10 15

Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu Arg Lys Lys Arg Phe 20 25 30

Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr Pro Leu Gln Gly Arg
35 40 45

Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr Asn Ile Ile Leu Gln 50 55 60

Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp Ser Glu Val Pro Tyr 65 70 75 80

Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser Gln Leu Cys Lys 85 90 95 Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser Pro Pro Pro Tyr Pro 105

Ser Val Pro Pro Pro Thr Pro Ser Ser Thr Asn Lys Asp Pro Pro Leu 115

Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly Val Asn Asn Glu Pro 130 135 140

Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro Leu Gln Ala Val Arg Gly 145 150 155 160

Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro Phe Ser Leu Ser Asp 165 170 175

Leu Lys Gln Ile Lys Ile Asp Leu Gly Lys Phe Ser Asp Asn Pro Asp 180 185 190

Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln Ser Phe Asp Leu Thr 195 200 205

Trp Arg Asp Ile Met Leu Leu Leu Asn Gln Thr Leu Thr Pro Asn Glu 210 215 220

Arg Ser Ala Ala Val Thr Ala Ala Arg Glu Phe Gly Asp Leu Trp Tyr 225 230 235 240

Leu Ser Gln Val Asn Asn Arg Met Thr Thr Glu Glu Arg Thr Thr Pro $245 \hspace{1.5cm} 250 \hspace{1.5cm} 255 \hspace{1.5cm}$

Thr Gly Gln Gln Ala Val Pro Ser Val Asp Pro His Trp Asp Thr Glu 260 265 270

Ser Glu His Gly Asp Trp Cys His Lys His Leu Leu Thr Cys Val Leu 275 280 285

Glu Gly Leu Arg Lys Thr Arg Lys Lys Pro Met Asn Tyr Ser Met Met 290 295 300

Ser Thr Ile Thr Gln Gly Lys Glu Glu Asn Pro Thr Ala Phe Leu Asp 305 310 315 320

ctttttcag atgggaaacg

Arg Le	u Arg	Glu	Ala 325	Leu	Arg	Lys	His	Thr 330	Ser	Leu	Ser	Pro	Asp 335	Ser	
Ile Gl	u Gly	Gln 340	Leu	Ile	Leu	Lys	Asp 345	Lys	Phe	Ile	Thr	Gln 350	Ser	Ala	
Ala As	p Ile 355	Arg	Lys	Asn	Phe	Lys 360	Ser	Leu	Pro						
<210>	32														
<211>	20														
<212>	DNA														
<213>	Homo	sapi	iens												
<400>	32	~+ ~ ~ ^	~+ ~ ~ +												2.0
atccaa	agig (grgac	JLaai	_d											20
<210>	33														
<211>	20														
<212>	DNA														
<213>	Homo	sap	iens												
<400>	33														

20

PCT09

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/869,927
DATE: 12/19/2001 (5)

Input Set : A:\ES.txt

Output Set: N:\CRF3\12192001\1869927.raw



```
7 <110> APPLICANT: PARANHOS-BACCALA, Glaucia
     10
             MALLET, Francois
     13
             VOISSET, Cecile
     19 <120> TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
             AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
     26 <130> FILE REFERENCE: 110048
     32 <140> CURRENT APPLICATION NUMBER: 09/869,927
C--> 35 <141> CURRENT FILING DATE: 2001-10-22
     41 <150> PRIOR APPLICATION NUMBER: PCT/FR00/00144
   44 <151> PRIOR FILING DATE: 2000-01-21
   50 <150> PRIOR APPLICATION NUMBER: FR 99/00888
   53 <151> PRIOR FILING DATE: 1999-01-21
  759 <160> NUMBER OF SEQ ID NOS: 33
  65 <170> SOFTWARE: PatentIn version 3.1
     71 <210> SEQ ID NO: 1
     74 <211> LENGTH: 1511
  77 <212> TYPE: DNA
  80 <213> ORGANISM: Homo sapiens

■ 86 <400> SEQUENCE: 1

   🔲 88 cctagaacgt attctggaga attgggacca atgtgacact cagacgctaa gaaagaaacg
                                                                               60
   🗐 92 atttatatte ttetgeagta eegeetggee acaatateet etteaaggga gagaaaeetg
                                                                              120
   🛀 96 getteetgag ggaagtataa attataacat catettacaq etagacetet tetqtagaaa
                                                                              180
   _{\sim}100 ggagggcaaa tggagtgaag tgccatatgt gcaaactttc ttttcattaa gagacaactc
                                                                               240
   104 acaattatgt aaaaagtgtg gtttatgccc tacaggaagc cctcagagtc cacctcccta
                                                                               300
     108 ccccagegte eceteceega etectteete aactaataag gacceeett taacceaaae
                                                                               360
    112 ggtccaaaag gagatagaca aaggggtaaa caatgaacca aagagtgcca atattccccg
                                                                               420
    116 attatgcccc ctccaagcag tgagaggagg agaattcggc ccagccagag tgcctqtacc
                                                                               480
    120 tttttctctc tcagacttaa agcaaattaa aatagaccta ggtaaattct cagataaccc
                                                                               540
    124 tqacqqctat attgatqttt tacaaqqqtt aggacaatcc tttgatctga catqqaqaqa
                                                                               600
    128 tataatgtta ctactaaatc agacactaac cccaaatgag agaagtgccg ctgtaactgc
                                                                               660
    132 agcccgagag tttggcgatc tttggtatct cagtcaggcc aacaatagga tgacaacaga
                                                                               720
    136 ggaaagaaca actcccacag gccagcaggc agttcccagt gtagaccctc attgggacac
                                                                               780
    140 agaatcagaa catggagatt ggtgccacaa acatttgcta acttgcgtgc tagaaggact
                                                                               840
    144 gaggaaaact aggaagaagc ctatgaatta ctcaatgatg tccactataa cacagggaaa
                                                                               900
    148 ggaagaaaat cttactgctt ttctggacag actaagggag gcattgagga agcatacctc
                                                                               960
    152 cctgtcacct gactctattg aaggccaact aatcttaaag gataagttta tcactcagtc
                                                                              1020
    156 agctgcagac attagaaaaa acttcaaaag tctgccttag gcccggagca gaacttagaa
                                                                              1080
    160 accetattta acttggcate etcagttttt tataatagag atcaggagga geaggegaaa
                                                                              1140
    164 cgggacaaac gggataaaaa aaaaaggggg ggtccactac tttagtcatg gccctcaggc
                                                                              1200
    168 aagcagactt tggaggctct gcaaaaggga aaagctgggc aaatcaaatg cctaataqqq
                                                                              1260
    172 ctggcttcca gtgcggtcta caaggacact ttaaaaaaga ttatccaagt agaaataagc
                                                                              1320
    176 cgcccccttg tccatgcccc ttacgtcaag ggaatcactg gaaggcccac tgccccaggg
                                                                              1380
    180 gatgaagata ctctgagtca gaagccatta accagatgat ccagcagcag gactgagggt
                                                                              1440
    184 gcccggggcg agcgccagcc catgccatca ccctcacaga gccccgggta tqtttqacca
                                                                              1500
    188 ttgagagcca a
                                                                              1511
    194 <210> SEQ ID NO: 2
    197 <211> LENGTH: 2009
```

RAW SEQUENCE LISTINGPATENT APPLICATION: US/09/869,927

DATE: 12/19/2001

TIME: 16:12:55

Input Set : A:\ES.txt

Output Set: N:\CRF3\12192001\1869927.raw

```
200 <212> TYPE: DNA
     203 <213> ORGANISM: Homo sapiens
     209 <220> FEATURE:
     212 <221> NAME/KEY: misc_feature
     215 <222> LOCATION: (1940)..(1940)
     218 <223> OTHER INFORMATION: n = a or g or c or t/u
     225 <400> SEQUENCE: 2
     227 atacgactac tatagggcga attgggccct ctagatgcat gctcgagcgg ccgccagtgt
                                                                                 60
     231 gatggatate tgeagaatte geeetttgte egetgtgete etgateeage gaggegeeea
                                                                                120
     235 ttgctgctcc caattgggct aaaggcttgc cattgttccc acacggctaa gtgcccgggt
                                                                                180
     239 tcatcctaat tgagctgaac actagtcact gggttccatg gttctcttcc atgacccacg
                                                                                240
     243 gcttctaata gagctctaat actcaccaca tggcccaaga ttccattcct tggaatccgt
                                                                                300
     247 gaggccaaga accccaggtc agagaacacg aggcttgcca ccgtcttgga agtggcccgc
                                                                                360
   251 cgccatettg ggagetetgg gageaaggae ceeecagtaa cattttggea accaeaaagg
                                                                                420
     255 gacctccaaa gcgatgggaa acattccccc caaggcaaaa acgcccctaa gatgtattct
                                                                                480
     259 ggagaattgg gaccaatgtg acactcagac gctaagaaag aaacgattta tattcttctg
                                                                                540
   263 cagtacegec tggccacaat atcetettea agggagagaa acetggette etgagggaag
                                                                                600
   267 tataaattat aacatcatct tacagctaga cctcttctgt agaaaggagg gcaaatggag
                                                                                660
   271 tgaagtgcca tatgtgcaaa ctttcttttc attaagagac aactcacaat tatgtaaaaa
                                                                                720
   275 gtgtggttta tgccctacag gaagccctca gagtccacct ccctacccca gcgtcccccc
                                                                                780
   1279 cccgactcct tecteaacta ataaggacee ceetttaace caaacggtee aaaaggagat
                                                                                840
   🐛 283 agacaaaggg gtaaacaatg aaccaaagag tgccaatatt ccccgattat gccccctcca
                                                                                900
   287 agcagtgaga ggaggagaat teggeecage cagagtgeet gtaeettttt eteteteaga
                                                                                960
   291 cttaaagcaa attaaaatag acctaggtaa attctcagat aaccctgacg gctatattga
                                                                               1020
     295 tgttttacaa gggttaggac aatcetttga tetgacatgg agagatataa tgttactact
                                                                               1080
   299 Lytttacaa yyyttayyac uutoottoga billya ja 299 aaatcagaca ctaaccccaa atgagagaag tgccgctgta actgcagccc gagagtttgg
                                                                              1140
   303 cgatctttgg tatctcagtc aggtcaacaa taggatgaca acagaggaaa gaacaactcc
                                                                              1200
   307 cacaggccag caggcagttc ccagtgtaga ccctcattgg gacacagaat cagaacatgg
                                                                               1260
   311 agattggtgc cacaaacatt tgctaacttg cgtgctagaa ggactgagga aaactaggaa
                                                                               1320
   =315 gaagectatg aattactcaa tgatgtccac tataacacag ggaaaggaag aaaatectac
                                                                              1380
     319 tgcttttctg gacagactaa gggaggcatt gaggaagcat acctccctgt cacctgactc
                                                                               1440
     323 tattgaaggc caactaatct taaaggataa gtttatcact cagtcagctg cagacattag
     327 aaaaaacttc aaaagtctgc cttaggcccg gagcagaact tagaaaccct atttaacttq
                                                                              1560
     331 gcatcctcag ttttttataa tagagatcag gaggagcagg cgaaacggga caaacgggat
                                                                              1620
     335 aaaaaaaaa gggggggtcc actactttag tcatggccct caggcaagca gactttggag
                                                                              1680
     339 gctctggaaa agggaaaagc tgggcaaatc aaatgcctaa tagggctggc ttccagtgcg
                                                                              1740
     343 gtctacaagg acactttaaa aaagattatc caagtagaaa taagccgccc ccttgtccat
                                                                              1800
     347 gccccttacg tcaagggaat cactggaagg cccactgccc caggggatga agatactctg
                                                                              1860
     351 agtcagaagc cattaaccag atgatccagc agcaggactg agggtgcccg gggcgagcgc
                                                                              1920
W--> 355 cageccatge cateaccetn acagageece gggtatgett gaccattgag agecaggagg
                                                                              1980
     359 ttaactgtct cctggacact ggcgcagcc
                                                                              2009
     365 <210> SEQ ID NO: 3
     368 <211> LENGTH: 1056
     371 <212> TYPE: DNA
     374 <213> ORGANISM: Homo sapiens
     380 <400> SEQUENCE: 3
     382 ctagaacgta ttctggagaa ttgggaccaa tgtgacactc agacgctaag aaagaaacga
                                                                                 60
     386 tttatattct tctgcagtac cgcctggcca caatatcctc ttcaagggag agaaacctgg
                                                                                120
     390 cttcctgagg qaagtataaa ttataacatc atcttacagc tagacctctt ctgtagaaag
                                                                                180
```

RAW SEQUENCE LISTING DATE: 12/19/2001 PATENT APPLICATION: US/09/869,927 TIME: 16:12:55

Input Set : A:\ES.txt

Output Set: N:\CRF3\12192001\I869927.raw

	•	
394	gagggcaaat ggagtgaagt gccatatgtg caaactttct tttcattaag agacaa	actca 240
398	3 caattatgta aaaagtgtgg tttatgeeet acaggaagee etcagagtee aeetee	ectac 300
402	cccagegtee ecteccegae tectteetea actaataagg acceeettt aaccea	aacg 360
406	5 gtccaaaagg agatagacaa aggggtaaac aatgaaccaa agagtgccaa tattcc	ccga 420
410) ttatgccccc tccaagcagt gagaggagga gaattcggcc cagccagagt gcctgt	tacct 480
414	l ttttctctct cagacttaaa gcaaattaaa atagacctag gtaaattctc agataa	accct 540
418	B gacggetata tigatgitti acaagggita ggacaateet tigateigae aiggag	gagat 600
422	2 ataatgttac tactaaatca gacactaacc ccaaatgaga gaagtgccgc tgtaac	etgca 660
	gcccgagagt ttggcgatct ttggtatctc agtcaggcca acaataggat gacaac	
430) gaaagaacaa ctcccacagg ccagcaggca gttcccagtg tagaccctca ttggga	icaca 780
434	gaatcagaac atggagattg gtgccacaaa catttgctaa cttgcgtgct agaagg	gactg 840
438	B aggaaaacta ggaagaagce tatgaattae teaatgatgt ceactataae acaggg	gaaag 900
	gaagaaaatc ttactgcttt tctggacaga ctaagggagg cattgaggaa gcatac	
	ctgtcacctg actctattga aggccaacta atcttaaagg ataagtttat cactca	
	gctgcagaca ttagaaaaaa cttcaaaagt ctgcct	1056
456	5 <210> SEQ ID NO: 4	
459) <211> LENGTH · 25	
462	2 <212 TYPE: DNA	
465	5 <213> ORGANISM: Homo sapiens	
471	5 <213> ORGANISM: Homo sapiens - <400> SEQUENCE: 4	
473	Coggacateca aagtgatggg aaacg	25
479	0 <210> SEQ ID NO: 5	
482	2 <211> LENGTH: 26	
485	5 <212> TYPE: DNA	
1 488	3 <213> ORGANISM: Homo sapiens	
494	<400> SEQUENCE: 5	
496	ggacaggaaa gtaagactga gaaggc	26
502	2 <210> SEQ ID NO: 6	
505	5 <211> LENGTH: 26	
	3 <212> TYPE: DNA	
	<213> ORGANISM: Homo sapiens	
	<pre><400> SEQUENCE: 6</pre>	
	cctagaacgt attctggaga attggg	26
	6 <210> SEQ ID NO: 7	
	3 <211> LENGTH: 26	
	<212> TYPE: DNA	
	<213> ORGANISM: Homo sapiens	
	<400> SEQUENCE: 7	
	tggctctcaa tggtcaaaca tacccg	26
	3 <210> SEQ ID NO: 8	
	<211> LENGTH: 21	
	<212> TYPE: DNA	
	<213> ORGANISM: Homo sapiens	
	<400> SEQUENCE: 8	_
	atgggaaacg ttccccccga g	21
	<210> SEQ ID NO: 9	
	<211> LENGTH: 21	
	<pre><212> TYPE: DNA </pre>	
280	<213> ORGANISM: Homo sapiens	

RAW SEQUENCE LISTING DATE: 12/19/2001 PATENT APPLICATION: US/09/869,927 TIME: 16:12:55

Input Set : A:\ES.txt

Output Set: N:\CRF3\12192001\1869927.raw

```
586 <400> SEQUENCE: 9
  588 ggcctaaggc agacttttga a
                                                                            21
  594 <210> SEQ ID NO: 10
  597 <211> LENGTH: 409
  600 <212> TYPE: PRT
  603 <213> ORGANISM: Homo sapiens
  609 <400> SEQUENCE: 10
  613 Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro
                                          10
 621 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
                 20
                                      25
 629 Ile Met Gly Asn Ile Pro Pro Lys Ala Lys Thr Pro Leu Arg Cys Ile
            35
                                  4.0
\square637 Leu Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu Arg Lys Lys Arg
                              55
4645 Phe Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr Pro Leu Gln Gly
647 65
                         70
                                              75
653 Arg Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr Asn Ile Ile Leu
661 Gln Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp Ser Glu Val Pro
 663
                 100
                                      105
669 Tyr Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser Gln Leu Cys Lys
             115
                                  120
                                                      125
677 Lys Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser Pro Pro Pro Tyr
□679 130
                              135
4685 Pro Ser Val Pro Pro Pro Thr Pro Ser Ser Thr Asn Lys Asp Pro Pro
                         150
                                              155
693 Leu Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly Val Asn Asn Glu
<u>1</u>695
                     165
                                          170
 701 Pro Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro Leu Gln Ala Val Arg
                 180
                                      185
 709 Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro Phe Ser Leu Ser
             195
                                  200
 717 Asp Leu Lys Gln Ile Lys Ile Asp Leu Gly Lys Phe Ser Asp Asn Pro
                              215
                                                  220
 725 Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln Ser Phe Asp Leu
                          230
                                              235
 733 Thr Trp Arg Asp Ile Met Leu Leu Leu Asn Gln Thr Leu Thr Pro Asn
                     245
                                          250
 741 Glu Arg Ser Ala Ala Val Thr Ala Ala Arg Glu Phe Gly Asp Leu Trp
                                      265
 749 Tyr Leu Ser Gln Val Asn Asn Arg Met Thr Thr Glu Glu Arg Thr Thr
             275
                                  280
 757 Pro Thr Gly Gln Gln Ala Val Pro Ser Val Asp Pro His Trp Asp Thr
         290
                              295
 765 Glu Ser Glu His Gly Asp Trp Cys His Lys His Leu Leu Thr Cys Val
                         310
                                             315
 773 Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Pro Met Asn Tyr Ser Met
 775
```

RAW SEQUENCE LISTING DATE: 12/19/2001 PATENT APPLICATION: US/09/869,927 TIME: 16:12:55

Input Set : A:\ES.txt

Output Set: N:\CRF3\12192001\1869927.raw

```
781 Met Ser Thr Ile Thr Gln Gly Lys Glu Glu Asn Pro Thr Ala Phe Leu
                  340
                                       345
                                                            350
 789 Asp Arg Leu Arg Glu Ala Leu Arg Lys His Thr Ser Leu Ser Pro Asp
              355
                                   360
 797 Ser Ile Glu Gly Gln Leu Ile Leu Lys Asp Lys Phe Ile Thr Gln Ser
                               375
 805 Ala Ala Asp Ile Arg Lys Asn Phe Lys Ser Leu Pro Lys Leu Ala Ala
                           390
                                               395
 813 Ala Leu Glu His His His His His His
 821 <210> SEQ ID NO: 11
 824 <211> LENGTH: 393
=827 <212> TYPE: PRT
830 <213> ORGANISM: Homo sapiens
__836 <400> SEQUENCE: 11
 840 Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Ile Met Gly Asn
 842 1
848 Ile Pro Pro Lys Ala Lys Thr Pro Leu Arg Cys Ile Leu Glu Arg Ile
<sup>4</sup>850
                  20
19856 Leu Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu Arg Lys Lys Arg
_{\scriptscriptstyle \pm} 864 Phe Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr Pro Leu Gln Gly
866
872 Arg Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr Asn Ile Ile Leu
<u>-</u>874 65
                          70
 880 Gln Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp Ser Glu Val Pro
                                           90
888 Tyr Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser Gln Leu Cys Lys
                  100
                                       105
 896 Lys Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser Pro Pro Pro Tyr
             115
                                   120
                                                        125
 904 Pro Ser Val Pro Ser Pro Thr Pro Ser Ser Thr Asn Lys Asp Pro Pro
         130
                              135
 912 Leu Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly Val Asn Asn Glu
                          150
                                               155
 920 Pro Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro Leu Gln Ala Val Arg
                                           170
 928 Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro Phe Ser Leu Ser
                  180
                                       185
 936 Asp Leu Lys Gln Ile Lys Ile Asp Leu Gly Lys Phe Ser Asp Asn Pro
 938
             195
                                   200
 944 Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln Ser Phe Asp Leu
                              215
                                                   220
 952 Thr Trp Arg Asp Ile Met Leu Leu Leu Asn Gln Thr Leu Thr Pro Asn
                          230
                                               235
 960 Glu Arg Ser Ala Ala Val Thr Ala Ala Arg Glu Phe Gly Asp Leu Trp
                      245
                                           250
 968 Tyr Leu Ser Gln Ala Asn Asn Arg Met Thr Thr Glu Glu Arg Thr Thr
                  260
                                       265
```



Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARYPATENT APPLICATION: **US/09/869,927**DATE: 12/19/2001 TIME: 16:12:56

Input Set : A:\ES.txt

Output Set: N:\CRF3\12192001\1869927.raw

```
L:35 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:355 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:1247 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18
L:1751 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1771 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1846 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:2033 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24
L:2041 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24
L:2045 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24
L:2502 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27
L:2952 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29
L:2956 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29
L:3012 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29
L:3024 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29
L:3227 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:3255 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:3299 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:3303 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:3359 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:3321 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:3467 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 L:3487 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:3499 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
    IJ
   D
```

09/869927
JC18 Rec'd PCT/PTO 0 9 JUL 2001
PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of

Glaucia PARANHOS-BACCALA, Francois MALLET, Cecile VOISSET

Application No.: US National Stage of PCT/FR00/00144

Filed: July 9, 2001

Docket No.: 110048

ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN

AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT

PRELIMINARY AMENDMENT

Director of the U.S. Patent and Trademark Office Washington, D. C. 20231

Sir:

For:

Prior to initial examination, please amend the above-identified application as follows:

IN THE CLAIMS:

Please replace claims 7, 11, 16 and 17 as follows:

- 7. (Amended) Transcription product which can be obtained by transcription of at least said portion of the *gag* gene of a fragment according to claim 1.
- 11. (Amended) Method according to claim 8, characterized in that the biological sample is a biological fluid chosen from serum, plasma, synovial fluid and urine.
- 16. (Amended) Reagent for detecting, in a biological sample, an autoimmune disease or monitoring pregnancy, comprising at least one fragment according to claim 1.
- 17. (Amended) Use of a fragment according to claim 1, for detecting, in a biological sample, susceptibility to an autoimmune disease, or monitoring pregnancy.

REMARKS

Claims 1-18 are pending. Claims 7, 11, 16 and 17 are amended to eliminate multiple dependencies. Prompt and favorable consideration on the merits is respectfully requested.

The attached Appendix includes marked-up copies of each rewritten claim (37 C.F.R. §1.121(c)(1)(ii)).

Respectfully submitted,

William P. Berridge Registration No. 30,024

Joel S. Armstrong Registration No. 36,430

WPB:JSA/zmc

Attached: APPENDIX

Date: July 9, 2001

OLIFF & BERRIDGE, PLC P.O. Box 19928 Alexandria, Virginia 22320 Telephone: (703) 836-6400 DEPOSIT ACCOUNT USE
AUTHORIZATION
Please grant any extension
necessary for entry;
Charge any fee due to our

Deposit Account No. 15-0461

APPENDIX

Changes to Claims:

The following are marked-up versions of the amended claims:

- 7. (Amended) Transcription product which can be obtained by transcription of at least said portion of the *gag* gene of a fragment according to <u>claim 1</u> any one of claims 1 to 6.
- 11. (Amended) Method according to claim 8 any one of claims 8 to 10, characterized in that the biological sample is a biological fluid chosen from serum, plasma, synovial fluid and urine.
- 16. (Amended) Reagent for detecting, in a biological sample, an autoimmune disease or monitoring pregnancy, comprising at least one fragment according to <u>claim 1.any</u> one of claims 1 to 6, one transcription/translation product, as obtained according to the method of claim 11, one sythentic peptide belonging to SEQ ID NO. 31 or one protein according to claim 15 or 16.
- 17. (Amended) Use of a fragment according to claim 1 any one of claims 1 to 6, of a transcription/translation product, as obtained according to the method of claim 11, one sythentic peptide belonging to SEQ ID NO. 31 or one protein according to claim 15 or 16, for detecting, in a biological sample, susceptibility to an autoimmune disease, or monitoring pregnancy.

10

20

25

30

35

5/PR1S

ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT

The present invention relates to an endogenous nucleic acid fragment of the retroviral type, integrated into the DNA of the human genome.

Retroviruses are RNA viruses which replicate through a process termed reverse transcription, mediated by an RNA-dependent DNA polymerase named reverse transcriptase (RT), which is encoded by the pol gene. The retroviral RNA also comprises at least two additional genes, which are the gag and env genes. The gag gene encodes the proteins of the backbone, i.e. the matrix, the capsid and the nucleocapsid. The env gene encodes the envelope proteins. The transcription is regulated by promoter regions located in the LTRs (Long Terminal Repeat) which border the 5'- and 3'-terminal ends of the retroviral genome.

In the course of evolution, humans or their ancestors have integrated material of retroviral origin into their genome subsequent to an infection. Specifically, when a cell is infected, the reverse transcriptase makes a DNA copy of the retroviral RNA, and this DNA copy may then possibly integrate into the human genome. Retroviruses can infect germinal cells and thus be transmitted to future generations by vertical Mendelien transmission. They are then referred to as endogenous retroviruses which are present in the form of proviral DNA integrated into the genome of all human cells. Most endogenous retroviruses are silent or defective. However, some of them have been able to conserve all or part of their initial properties and may be activated under specific conditions. The expression of endogenous retroviruses can range from the transcription of viral genes to the production of viral particles.

15

20

These endogenous retroviruses may be associated directly or indirectly with the development of certain pathological conditions.

5 Endogenous retroviral structures may be in a complete LTR-gag-pol-env-LTR form or in truncated forms.

previous patent application in Thus, (PCT/FR98/01442), the applicant screened a cDNA library using a Ppol-MSRV probe (SEQ ID NO. 18) and detected overlapping clones which allowed it to reconstruct a putative genomic RNA of 7582 nucleotides. This genomic RNA has an R-U5-gag-pol-env-U3-R structure. A "blastn" interrogation over several databases using the reconstructed genome made it possible to show that there is a considerable amount of related genomic (DNA) sequences in the human genome, which are found on several chromosomes. Thus, the applicant demonstrated the existence of partial structures of the retroviral type in the human genome and envisaged their potential role in the development of autoimmune diseases, in unsuccessful pregnancy or pathological conditions of pregnancy.

Autoimmune diseases which may be mentioned by way of example are multiple sclerosis, rhumatoid arthritis, lupus erythematosus disseminatus, insulin-dependent diabetes and/or pathologies which are associated with them.

30

35

The isolation and sequencing of overlapping cDNA fragments and the identification of genomic (DNA) clones corresponding to the isolated DNA clones, described in the applicant's abovementioned PCT patent application, are incorporated herein by way of reference.

Isolation and sequencing of overlapping cDNA fragments:

The information regarding the organization of the novel endogenous retroviruses named, by the family of applicant, HERV-W was obtained by testing a placenta cDNA library (Clontech cat#HL5014a) with the Ppol-MSRV (SEO ID NO. 18) and Penv-C15 (SEQ ID NO. 19) probes and then carrying out a "gene walking" technique using the novel sequences obtained. The experiments were carried out with reference to the recommendations of the supplier of the library. PCR amplifications on DNA were also used in order to understand this organization.

The following clones were selected and sequenced:

- Clone cl.6A2 (SEQ ID NO. 20): 5' untranslated region 15 of HERV-W and a portion of gag.
 - Clone cl.6A1 (SEQ ID NO. 21): gag and a portion of pol.
 - Clone cl.7A16 (SEQ ID NO. 22): 3' region of pol.
- Clone cl.Pi22 (SEQ ID NO. 23): 3' region of pol and 20 start of env.
 - Clone cl.24.4 (SEQ ID NO. 24): spliced RNA comprising a portion of the 5' untranslated region of HERV-W, the end of pol and the 5' region of env.
- Clone cl.C4C5 (SEQ ID NO. 25): end of env and 3' 25 untranslated region of HERV-W.
 - Clone cl.PH74 (SEQ ID NO. 26): subgenomic RNA: untranslated region of HERV-W, end of pol, env, and 3' untranslated region of HERV-W.
- Clone cl.PH7 (SEQ ID NO. 27): multispliced RNA: 5' 30 untranslated region of HERV-W, end of env untranslated region of HERV-W.
 - Clone cl.Pi5T (SEQ ID NO. 28): partial pol gene and U3-R region.
- Clone cl.44.4 (SEQ ID NO. 29): R-U5 region, gag gene 35 and partial pol gene.

A total sequence model for HERV-W was produced with the aid of these clones, by carrying out sequence alignments. The spliced RNAs were revealed and also the potential splice donor and acceptor sites. The LTR, gag, pol and env entities were defined by studying similarity with existing retroviruses.

The putative genetic organization of HERV-W in the RNA form is as follows (SEQ ID NO. 30):

gene 1..7582.

Location of the clones on the reconstructed genomic RNA sequence:

15 cl.6A2 (1321 bp) 1-1325;

cl.PH74 (535+2229= 2764 bp) 72-606 and 5353-7582;

cl.24.4 (491+1457= 1948 bp); 115-606 and 5353-6810;

cl.44.4 (2372 bp) 115-2496;

cl.PH7 (369+297= 666 bp) 237-606 and 7017-7313; cl.6A1

20 (2938 bp) 586-3559;

cl.Pi5T (2785+566= 3351 bp) 2747-5557 and 7017-7582;

cl.7A16 (1422 bp) 2908-4337;

cl.Pi22 (317+1689= 2006 bp) 3957-4273 and 4476-6168;

cl.C4C5 (1116 bp) 6467-7582

25 5'LTR 1..120

/note="R of 5'LTR (5' end uncertain [sic]"

121..575

/note="U5 of 5'LTR"

misc. 579..596

/note="PBS, primer binding site, for tRNA-W"

misc. 606

/note="splice junction (splice donor site ATCCAAAGTG-GTGAGTAATA and splice acceptor site CTTTTTCAG-ATGGGAAACG, clone RG083M05,

35 GenBank accession AC000064)"

misc. 5353

/note="splice acceptor site for ORF1 (env)"

misc. 5560

30

/note="splice donor site"

ORF 5581..7194

/note="ORF1 env 538 AA"

/product-="envelope"

5 misc. 7017

/note="splice acceptor site for ORF2 and
ORF3"

ORF 7039..7194

/note="ORF2 52 AA"

10 ORF 7112..7255

/note="ORF3 48 AA"

misc. 7244..7254

/note="PPT, polypurine tract"

3'LTR 7256..7582

15 /note="U3-R of 3' LTR (U3-R junction

undetermined)

misc. 7563..7569

polyadenylation signal

20 Identification of genomic (DNA) clones corresponding to the isolated DNA clones:

A "blastn" interrogation over several databases, using the reconstructed genome, showed that there is a considerable amount of related sequences in the human genome. Approximately 400 sequences were identified in GenBank and more than 200 sequences in the EST bank, most of them in the antisense orientation. The 4 most significant sequences in terms of size and similarity are the sequences of the following genomic (DNA) clones:

Human clone RG083M05 (gb AC000064), the chromosomal location of which is 7q21-7q22,

Human clone BAC378 (gb U85196, gb AE000660)

35 corresponding to the alpha/delta locus of the T-cell receptor, located at 14q11-12,

Human cosmid Q11M15 (gb AF045450) corresponding to region 21q22.3 of chromosome 21,

10

15

20

25

30

35

Cosmid U134E6 (embl Z83850) on chromosome Xq22.

The location of the aligned regions for each of the clones is indicated and the chromosome to which they belong is indicated between square brackets (Figure 6 [sic]). The percentage similarity (without the large deletions) between the 4 sequences and the reconstructed genomic RNA is indicated, and also the presence of repeat sequences at each end of the genome and the size of the longest open reading frames (ORFs). Repeat sequences were found at the ends of 3 of these The reconstructed sequence is contained within clone RG083M05 (9.6 Kb) and exhibits 96% similarity. However, clone RG083M05 has a 2 Kb insertion located immediately downstream of the untranslated region (5' UTR). This insertion is also found in two other genomic clones which have a 2.3 Kb deletion immediately upstream of the 3' untranslated region (3' UTR). No clone contained the three functional gag, pol and env open reading frames (ORFs). Clone RG083M05 shows a 538 amino acid (AA) corresponding to a whole envelope. Cosmid Q11M15 contains two major contiguous ORFs of 413 AA (frame 0) and 305 AA (frame +1) corresponding to a truncated pol polyprotein.

An endogenous nucleic acid fragment has now been found and isolated, which is integrated into the DNA of the human genome and which comprises or consists of at least one portion of the gag gene of an endogenous retrovirus associated with an autoimmune disease, or with unsuccessful pregnancy or pathological conditions of pregnancy, this portion at least encoding, directly or indirectly, an expression product. Of course, the invention also comprises the sequence complementary to said fragment.

Advantageously, the fragment defined above also satisfies at least any one of the following characteristics:

5 It comprises, or consists of, said whole gag gene;

Said portion of the fragment at least encodes the matrix and the capsid;

10 It comprises, or consists of, SEQ ID NO. 1, SEQ ID NO. 2, SEQ ID NO. 3 or the sequence complementary to any one of said sequences;

It is located on at least one of human chromosomes 1, 3, 6, 7 and 16, it is preferably located on at least chromosome 3;

The product of expression of said portion is messenger RNA;

20

25

30

35

The product of expression of said portion is immunologically recognized by antibodies present in a biological sample from a patient suffering from an autoimmune disease, such as multiple sclerosis; preferably, the biological fluid is chosen from serum, plasma, synovial fluid and urine.

Another subject of the invention is an endogenous transcription product which is in isolated form and which can be obtained by transcription of at least said portion of the gag gene of a fragment of the invention.

The invention also relates to a method for detecting endogenous nucleotide sequences belonging to a fragment of the invention, comprising the following steps:

a prior step of extraction of the cellular DNA from a tissue or biological fluid is carried out, and then at

least one cycle of amplification of the cellular DNA is carried out, for instance by PCR, using primers in particular chosen from SEQ ID NO. 4 to SEQ ID NO. 9 and SEQ ID NO. 12 to SEQ ID NO. 17,

5

the cellular DNA present in the sample is brought into contact with a given probe which is capable of hybridizing with a fragment as defined above and of forming a hybridization complex, said probe comprising at least 15 contiguous nucleotides, preferably 17 and advantageously 19 contiguous nucleotides, of SEQ ID NO. 3, or consisting of SEQ ID NO. 3, under suitable conditions for the hybridization, in particular under conditions of high stringency, and

15

10

the hybridization complexes formed are detected by any suitable means.

20

Advantageously, the probe is labeled with a tracer, such as for example a radioactive tracer or an enzyme.

The invention also relates to a method for detecting endogenous nucleotide sequences belonging to a fragment of the invention, comprising the following steps:

25

a prior step of extraction of the cellular DNA from a tissue or biological fluid is carried out, and then at least one cycle of amplification of the cellular DNA is carried out, for instance by PCR, using primers in particular chosen from SEQ ID NO. 4 to SEQ ID NO. 9 and SEQ ID NO. 12 to SEQ ID NO. 17,

a step of in vitro transcription/translation of the amplified product is carried out, and

35

30

the product derived from the transcription/translation step is reacted with a serum or plasma from a patient with an autoimmune disease.

The invention also relates to a method for studying and/or monitoring T-cell proliferation in vitro, according to which the T cells from a patient are brought into contact with either transcription/translation products (SEQ ID NO. 31), as obtained according to the method above, or synthetic peptides derived from or belonging to SEQ ID NO. 31.

Another subject of the invention is a method for the in situ molecular labeling of chromosomes isolated from patients, in which a probe labeled with any suitable tracer, and comprising all or part of SEQ ID NO. 3, is used.

15

20

30

35

The invention also relates to:

a recombinant protein obtained using an expression cassette in a bacterial host, characterized in that its protein sequence consists of SEQ ID NO. 31; the bacterial host is in particular *E. coli*;

a reagent for detecting an autoimmune disease or monitoring pregnancy, comprising at least one fragment or one protein of the invention;

the use of a fragment or of a protein of the invention for detecting, in a biological sample, susceptibility to an autoimmune disease, or monitoring pregnancy; in particular, the autoimmune disease is multiple sclerosis.

Before setting out the present invention in greater detail, the definition of certain terms employed in the description and claims is given.

The expression "expression product" means any product derived from the retroviral DNA integrated into the

10

15

20

25

30

35

human genome, including the transcription products (messenger RNA) and the products derived from the translation of the messenger RNA obtained. In the latter case, and by way of example, the product may be a peptide or a protein which is functional or functionalizable, i.e. which can become functional.

The expression "portion encoding, directly or indirectly, an expression product" is intended to mean a portion which, by itself, comprises at least all or part of an open reading frame from which it is possible to deduce an amino acid sequence, and the coding capacity of which can be induced by elements such as, for example, those which may have promoter activity. This definition includes the variability which may be found in the coding nucleic acid sequence, provided that the above conditions are respected.

Example 1: Location of the gag gene of the HERV-W family on human chromosomes using the Southern blot technique

In order to locate the gag gene of the HERV-W family, a probe corresponding to this gene from MSRV was hybridized on a nylon membrane (Hybond® N+, Amersham) containing 5 µg of DNA from 24 somatic cell hybrids [human x rodents] (isolated human genomic DNA: 22 autosomal chromosomes and 2 sex chromosomes) and 3 control DNAs (human, mouse and hamster), digested with the EcoRI restriction enzyme.

The following probe is used: Pgag-C12 identified by SEQ ID NO. 3 corresponding to the coding region (of 1056 bp) of the clone MSRV gag C12.

1.1- Production of clone 2, Cl2, containing, in the 3' region, a portion homologous to the *pol* gene, corresponding to the protease gene, and a portion

homologous to the gag gene, corresponding to the nucleocapsid, and a 5' coding region, corresponding to the gag gene, more specifically the matrix and capsid of MSRV-1.

5

10

15

A PCR amplification was carried out on total RNA extracted from 100 μ l of plasma from a patient suffering from MS. A water control, treated under the same conditions, was used as a negative control. The cDNA synthesis was carried out with 300 pmol of a random primer (Gibco-BRL, France) and the "Expand RT" reverse transcriptase (Boehringer Mannheim, France), according to the conditions recommended by the company. A PCR (polymerase chain reaction) amplification was carried out with the Taq polymerase enzyme (Perkin Elmer, France) using 10 μ l of cDNA under the following conditions: 94°C 2 min, 55°C 1 min and 72°C 2 min, then 94°C 1 min, 55°C 1 min and 72°C 2 min for 30 cycles and 72°C for 7 min, with a final reaction volume of 50 μ l.

20

The primers used for the PCR amplification are as follows:

- 5' primer, identified by SEQ ID NO. 4
- 25 5' CGG ACA TCC AAA GTG ATG GGA AAC G 3';
 - 3' primer, identified by SEQ ID NO. 5
 - 5' GGA CAG GAA AGT AAG ACT GAG AAG GC 3'

A second "nested" PCR amplification was carried out with 5' and 3' primers located inside the region already amplified. This second PCR was carried out under the same experiment conditions as those used in the first PCR, using 10 μ l of the amplification product derived from the first PCR.

35

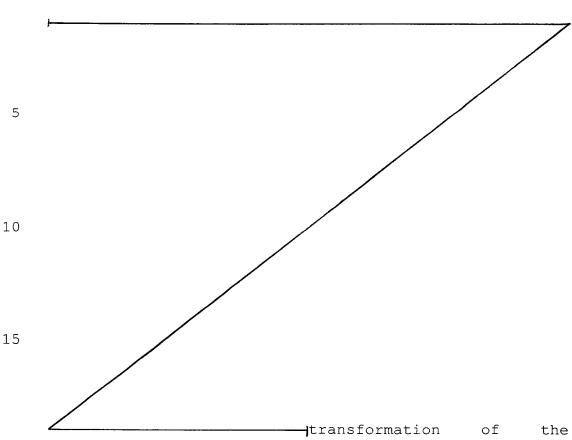
The primers used for the nested PCR amplification are as follows:

20

25

30

35



ligation in E. coli bacteria, the ligation mixture was plated out. At the end of the procedure, the white colonies of recombinant bacteria were picked in order to be cultured and to allow the extraction of the incorporated plasmids according to minipreparation" procedure (J. Sambrook, E.F. Fritsch and T. Maniatis, Molecular Cloning, a laboratory manual, Cold Spring Harbour Laboratory Press, 1989). The plasmid preparation from each recombinant colony was cleaved with the Eco RI restriction enzyme and analyzed on agarose gel. The plasmids possessing an insert which was detected under UV light after staining the gel with ethidium bromide were selected in order to sequence the insert after hybridization with a primer complementary to the T7 promoter present on the cloning plasmid from the TA Cloning Kit®. The reaction prior to the sequencing was then carried out according to the method recommended for using the "Prism® Ready Reaction

10

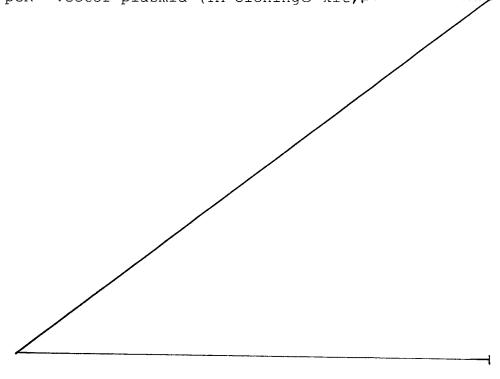
Amplitaq® FS, DyeDeoxy $^{\text{TM}}$ Terminator" sequencing kit (Applied Biosystems, ref. 402119) and the automatic sequencing was carried out on the Applied Biosystems 373 A and 377 machines, according to the manufacturer's instructions.

The clone obtained, named Cl2, makes it possible to define a 1511 bp region which has an open reading frame in the N-terminal region of 1089 bp (fragment 434-1521 of SEQ ID NO. 2) encoding 363 amino acids (SEQ ID NO. 31) corresponding to the matrix and capsid regions of the gag gene.

The nucleotide sequence of Cl2 is identified by SEQ ID NO. 1. It is represented in Figure 2 with the potential amino acid reading frames.

1.2- Production of the MSRV gag cl2 probe

The probe was obtained after PCR amplification, using the pCR $^{\text{TM}}$ vector plasmid (TA Cloning® kit,



AMENDED PAGE

out and labeled with $[\alpha-P^{32}]$ using random primers (Gibco-BRL, France) in accordance with instructions of the "Ready-to-go DNA labeling" kit (Pharmacia Biotech). The unincorporated nucleotides were removed with a G-50 Quick Spin column (Boehringer, Mannheim).

1.3- Southern blot

The hybridization conditions are as follows:

10

15

20

25

5

After prehybridization for 4 hours (in 5X SSC, 1X Denhardt's, 0.1% SDS, 50% formamide, 20 mM Tris-HCl, pH = 7.5, and 0.1 mg/ml of herring sperm DNA), the nylon membrane containing the human chromosomes was hybridized (in 5X SSC, 1X Denhardt's, 0.1% SDS, 50% formamide, 20 mM Tris-HCl, pH = 7.5, 0.1 mg/ml of herring sperm DNA and 5% dextran sulfate) for 18 hours at 42°C with the 32P-labeled 1056 bp gag cl2 DNA probe (SEQ ID NO. 3). After hybridization, the membrane (The BIOS Monochromosomal Somatic Cell Hybrid blot, from Quantum Bioprobe) hybridized with the gag probe was washed twice in 2X SSC/0.2% SDS solution for 15 min at room temperature, and twice (in 0.2% SSC/0.2% SDS) for 15 min at $45\,^{\circ}\text{C}$. After washing, the membrane was exposed to the X-ray film at $-80\,^{\circ}\text{C}$ in the presence of an amplifying screen.

The results are given in Table 1 hereinafter.

30 In this table:

m, which signifies mouse, and h, which signifies hamster, correspond to the recipient cells for the human chromosomal DNA.

35

The number indicated under each chromosome corresponds to the number of bands encountered.

The total number of copies of the gag gene is 66.

- 15 -

Table 1

			T
Ham-	ster		0
Mouse			0
X	ㅂ		0
×	ų		4
22	h		0
21	Ħ		4
20	ш		0
19	ч		1
18	Ч		2
17	Ħ		3
16	Ħ		0
15	ᅺ		3
14	Ч		1
13	<u>,</u>		3
12	h		9
11	h		3
10	Ч		4
Q	ч		2
8	٦		Э
7	ᅺ		2
9	ᅩ		ж
5	ч		2
4	<u>.</u> d		9
3	ч		9
2	ч		0
-	ш		5
Chromo No.	Rodent	parent	Gag probe

20

25

Example 2: PCR amplification of the gag gene of the HERV-W family each of on the isolated human chromosomes; verification of the specificity of the amplifications by Southern blot; "in transcription/translation (PTT) test using the PCR products, in order to verify the coding capacity and discover which of the human chromosomes have open reading frames for the gag gene of the HERV-W family.

10 2.1- PCR amplification

In order to amplify the HERV-W gag gene, a PCR was carried out on each isolated human chromosome [NIGMS human/rodent somatic cell hybrid panel #2. The human monochromosomal NIGMS somatic hybrid mapping panel #2, described by H.L. Drwinga et al. and B.L. Dubois et al., obtained from the Coriell Institute (Camden, NJ)] with the Taq polymerase enzyme (Perkin Elmer, France) using: 40 pmol of each primer, 25 mM of each dNTP (Pharmacia), 2.5 mM of MgCl₂, 2.5 U of Taq polymerase in the standard PCR buffer (Perkin Elmer) and 300 ng of isolated chromosome DNA, in a final volume of 100 μ l. The PCR conditions for amplifying the gag region are as follows: 3 min at 94°C; then 1 min at 94°C, 1 min at 55°C and 3 min at 72°C for 30 cycles, and 7 min at 72°C.

The primers used for the PCR amplification of the gag gene, from an ATG introduced into the HERV-W gag sequence on each isolated human chromosome are as follows:

- 5' primer, identified by SEQ ID NO. 14
5'-TTT GGT AAT ACG ACT CAC TAT AGG GCA GCC ACC ATG GGA
35 AAC GTT CCC CCC GAG-3'.

The primer contains the T7 RNA polymerase promoter sequence, a "spacer", the Kozak sequence (translation

initiation site in eukaryotes) and the 5' gag sequence starting from the HERV-W ATG.

- 3' primer, identified by SEQ ID NO. 15 5'-TTTTTTTTTTTTTTTTTTCAGGCTGCGCCAGTGTCCAGGAGAC-3'.

5

primer contains a poly-A tail (in order to stabilize the transcription of the RNA, represented by 18 T bases), a stop codon (represented by TCA) and the sequence of the MSRV-1 protease gene (G+E+A).

10

15

For the amplification of the HERV-W gag gene using oligonucleotides defined in the LTR and protease regions of HERV-W, with the Taq polymerase enzyme (Perkin Elmer, France), the PCR conditions were as follows: 3 min at 94°C; then 1 min at 94°C, 1 min at 60°C and 2 min at 72°C, 35 cycles; followed by 7 min at 72°C, with 50 ng of each monochromosomal DNA.

The primers used for the PCR amplification of the gag 20 gene using the oligonucleotide defined in the HERV-W LTR sequence, on each isolated human chromosome, are as follows:

- 5' primer, identified by SEQ ID NO. 16
- 5'-TGTCCGCTGTGCTCCTGATC-3'
- 3' primer, identified by SEQ ID NO. 17 25 5'-TTTTTTTTTTTTTTTTTTCAGGCTGCGCCAGTGTCCAGGAGAC-3'.

primer contains a poly-A tail (in order to stabilize transcription of the RNA, represented by 18 T 30 bases), a stop codon (represented by TCA) and the sequence of the MSRV-1 G+E+A protease gene.

The PCR amplifications were carried out in an MJ Research PTC200 Peltier Thermal cycler machine. The PCR 35 products (10 μ l of each PCR product) were analyzed in a gel of 1% agarose in 1X TBE (Tris-HCl, borate, EDTA). In order to verify the specificity of the amplification products, 3 μ l of each PCR product were analyzed in

agarose gel and then transferred onto a nylon membrane (Hybond® $-N^+$, Amersham) (Southern blot) using 0.4 N The hybridization with the gag cl2 probe (1056 bp) (J. Sambrook et al., 1989) was carried out under the following conditions: after prehybridization for 4 hours (in 5X SSC, 1X Denhardt's, 0.1% SDS, 50% formamide, 20 mM Tris-HCl, pH = 7.5, and 0.1 mg/ml of herring sperm DNA), the nylon membrane was hybridized (in 5X SSC, 1X Denhardt's, 0.1% SDS, 50% formamide. 20 mM Tris-HCl, pH = 7.5, 0.1 mg/ml of herring sperm DNA and 5% dextran sulfate), for 18 hours at 42°C with the ³²P-labeled gag DNA probe. The gag PCR products from each isolated human chromosome were washed once, for 15 min at room temperature, in a solution of 2X SSC, 0.2% SDS; twice, for 15 min each wash at 65°C, in a solution of 0.2X SSC, 0.1% SDS; twice, for 15 min each at 65°C, in a solution of 0.1% SSC, 0.1% SDS; and twice, for 30 min each at room temperature, in a solution of 0.1% SSC, 0.1% SDS.

20

10

15

Part of the remaining volume (4 μ l) of the PCR amplification products was used for the PTT "in vitro" transcription/translation test (Roest PAM et al., 1993) (Promega, France). The remaining volume was used for 25 the cloning in the pCR® 2.1-TOPO vector (Invitrogen) in accordance with the instructions with the kit, and for the sequencing with the method recommended for using the "PRISM" Ready Reaction Amplitaq® FS, DyeDeoxy TM Terminator" sequencing kit (Applied Biosystems, ref. 402119), and the automatic sequencing was carried 30 out on Applied Biosystems 373A and 377 machines, according to the manufacturer's instructions.

The portion encoded (SEQ ID NO. 31) by the 2009 bp fragment (SEQ ID NO. 2) was amplified by PCR with the Pwo enzyme (5 U/ μ l) (Boehringer Manneim, France) using 1 μ l of the minipreparation of the gag clone DNA (SEQ ID NO. 3) under the following conditions: 95°C 1 min,

60°C 1 min and 72°C 2 min for 25 cycles, with a final reaction volume of 50 μ l, using the primers:

- 5' primer (Bam HI) (SEQ ID NO. 8):
- 5' ATG GGA AAC GTT CCC CCC GAG 3' (21 mer), and
- 3' primer (Hind III), identified by SEQ ID NO. 9 5[sic] GGC CTA AGG CAG ACT TTT GAA 3' (21 mer).

The fragment obtained after PCR was linearized with Bam HI and Hind III and subcloned into the pET28C and 10 pET21C expression vectors (NOVAGEN) linearized with Bam HI and Hind III. The DNA of the 1089 bp fragment in the two expression vectors were sequenced according to the method recommended for using the "PRISM" Ready Amplitaq® FS, Reaction DyeDeoxyTM Terminator" 15 sequencing kit (Applied Biosystems, ref. 402119) and the automatic sequencing was carried out on Applied Biosystems 373A and 377 machines, according to the manufacturer's instructions.

20 The expression of the nucleotide sequence of the 1089 bp fragment of the gag clone by the pET28C and pET21C expression vectors is identified by SEQ ID NO. 10 and SEQ ID NO. 11, respectively.

25 2.2- "In vitro" transcription/translation test (PTT, Promega)

This test was carried out in order to pinpoint the human chromosomes which have open reading frames for 30 the gag gene of the HERV-W family.

mixture containing 12.5 μ l of TNT® rabbit reticulocyte lysate (Promega), 1 μl of TNT® reaction buffer (Promega), 0.5 μ l of TNT® RNA polymerase (Promega), 0.5 μ l of a 1 mM mixture of amino acids 35 minus methionine, 2 μ l of ³⁵S-methionine</sup> (1000 Ci/mmol) at 10 mCi/ μ l (Amersham), 0.5 μ l of RNasin® ribonuclease inhibitor at 40 U/ μ l, 4 μ l of PCR amplification

products (equivalent to 1 μ g) from each human chromosome and 4 μ l of water, in a reaction volume of 25 μ l, [lacuna]. This mixture was incubated at 30°C for 90 min.

5

The gag proteins corresponding to the products of transcription/translation of the gag gene of the HERV-W family from each human chromosome, amplified by PCR, were revealed by 10% polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate (SDS)-PAGE after exposure of the gel to the X-ray film at room temperature in the presence of an amplifying screen.

The results are given in Table 2 hereinafter.

15

10

In this table, the number indicated under each chromosome corresponds to the molecular mass (kDa) of the proteins visualized in polyacrylamide gel in the presence of SDS.

20

Table 2

Water							
Y		Ч					
×		'n			_		
22		ų					
21		ш					
20		ш					
19		h					
18		4					
17		ш					
16		ш		14			
15		Ч					
14		Ч					
13		h					
12		ч					
11		h					
10		Ч					
6		Ч					
ω		Ч					
7		q		22			
2 3 4 5 6		Ч		23	18	17	
2		ч					
4		ч					
က		┖		45	25	20	17
2		ч ч ч ч щ ш					
Н		Ħ		28	23	18	
Chromo Chromo	No.	Rodent	parent		gag		
Chromo	No.	Rodent	parent parent		PCR gag		•

Example 3: Expression of the gag clone in Escherichia coli, and reaction with human sera

The coding region SEQ ID NO. 2 was expressed in Escherichia coli, and then the products thus expressed were tested against serum from patients suffering from MS, and also against serum from healthy patients.

The constructs pET28c-gag clone (1089 bp) and pET21Cgag clone (1089 bp) synthesize, in the BL21 (DE3) 10 bacterial strain, an N-terminal and C-terminal fusion protein for the pET28C vector, and a C-terminal fusion protein for the pET21C vector, with 6 histidine residues and an apparent molecular mass approximately 45 kDa, which are revealed by SDS-PAGE 15 polyacrylamide gel electrophoresis (U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature, 1970, 227: 680-685).

20

25

The reactivity of the protein with respect to an antihistidine monoclonal antibody (DIANOVA) demonstrated using the Western blot technique (H. Towbin et al., Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications, Proc. Natl. Acad. Sci. USA, 1979, 76: 4350-4354).

The recombinant proteins pET28C-gag clone (1089 bp) and 30 pET21C-gag clone (1089 bp) were visualized, by SDS-PAGE, in the insoluble fraction after enzymatic digestion of the bacterial extracts with $50 \mu l$ of lysozyme (10 mg/ml) and lysis by ultrasound.

35 The antigenic properties of the recombinant antigens pET28C-gag clone (1089 bp)and pET21C-gag (1089 bp) were tested by Western blot solubilization of the bacterial pellet with 2% SDS and 50 mM β -mercaptoethanol. After incubation with the sera from patients suffering from multiple sclerosis, the sera from the neurological controls and the blood transfusion center (BTC) control sera, the immunocomplexes were detected using an alkaline phosphatase-coupled anti-human IgG and IgM goat serum.

The results are given in Table 3 hereinafter.

10 Table 3

Reactivity of sera from patients suffering from multiple sclerosis and controls, with the recombinant gag protein produced in $E.\ coli^a$

15

5

DISEASE	NUMBER OF INDIVIDUALS	NUMBER OF
	TESTED	INDIVIDUALS
MS	15	6
		2(+++), 2(++),
		2 (+)
NEUROLOGICAL CONTROLS	2	1 (+++)
HEALTHY CONTROLS (BTC)	22	1 (+/-)

(a) The strips containing 1.5 μg of recombinant gag antigen show reactivity against sera diluted 100-fold. The Western blot interpretation is based on the 20 presence or absence of a gag-specific band on the strips. Positive and negative controls are included in each experiment.

These results show that, under the technical conditions used, approximately 40% of the human multiple sclerosis sera tested react with the recombinant gag protein.

30

CLAIMS

- 25 -

- Nucleic acid fragment, characterized in that it 1. consists of at least one portion of the gag gene of an endogenous retrovirus associated with an 5 autoimmune disease, or with unsuccessful pregnancy or pathological conditions of pregnancy, said portion being chosen from SEQ ID NO. 2 and any series of contiguous nucleotides belonging to SEQ ID NO. 2 but not belonging to SEQ ID NO. 1 and 10 encoding an expression product, or the sequence complementary to said fragment.
- Fragment according to claim 1, characterized in 2. that it can be isolated from at least one of human 15 chromosomes 1, 3, 6, 7 and 16.
 - Fragment according to claim 2, characterized in 3. that it can be isolated from at least chromosome 3.
 - Fragment according to claim 1, characterized in 4. that the expression product is messenger RNA.
- Fragment according to claim 1, characterized in 25 5. that the expression product is immunologically recognized by antibodies present in a biological sample from a patient suffering from an autoimmune disease.
- Fragment according to claim 5, characterized in 6. that the autoimmune disease is multiple sclerosis.
- Transcription product which can be obtained by 7. 35 transcription of at least said portion of the gag gene of a fragment according to any one of claims 1 to 6.

20

25

35

- 8. Method for detecting, in a biological sample, nucleotide sequences which are integrated into the DNA of the human genome and which belong to the gag gene of an endogenous retrovirus associated with an autoimmune disease, or with unsuccessful pregnancy or pathological conditions of pregnancy, characterized in that:
- a prior step of extraction of the cellular DNA of said biological sample is carried out, and then at least one cycle of amplification of the cellular DNA is carried out, for instance by PCR, using primers in particular chosen from SEQ ID NO. 4 to SEQ ID NO. 9 and SEQ ID NO. 12 to SEQ ID NO. 17,
 - the cellular DNA present in the sample brought into contact with a given probe which is capable of hybridizing with a said nucleotide sequence and of forming a hybridization complex, said probe comprising at least 15 contiguous nucleotides, preferably 17 and advantageously 19 contiguous nucleotides, of SEQ ID NO. 3, or consisting of SEQ ID NO. 3, under suitable conditions for the hybridization, in particular under conditions of high stringency, and
- the hybridization complexes formed are detected by any suitable means.
 - 9. Method according to claim 8, characterized in that the probe is labeled with a tracer, such as for example a radioactive tracer or an enzyme.
 - 10. Method for detecting, in a biological sample, nucleotide sequences which are integrated into the

10

20

25

DNA of the human genome and which belong to the gag gene of an endogenous retrovirus associated with an autoimmune disease, or with unsuccessful pregnancy or pathological conditions of pregnancy, characterized in that:

- a prior step of extraction of the cellular DNA of said biological sample, optionally derived from isolated chromosomes, is carried out, and then at least one cycle of amplification of the cellular DNA is carried out, for instance by PCR, using primers in particular chosen from SEQ ID NO. 4 to SEQ ID NO. 9 and SEQ ID NO. 12 to SEQ ID NO. 17,
- a step of in vitro transcription/translation of the amplified product is carried out, and
 - the product derived from the transcription/translation step is reacted with a serum or plasma from a patient with an autoimmune disease.
 - 11. Method according to any one of claims 8 to 10, characterized in that the biological sample is a biological fluid chosen from serum, plasma, synovial fluid and urine.
- 12. Method for studying and/or monitoring T-cell proliferation in vitro, according to which the T cells from a patient are brought into contact with either transcription/translation products, as obtained according to the method of claim 11, or synthetic peptides belonging to SEQ ID NO. 31.
- 35 13. Method for the in situ molecular labeling of chromosomes isolated from patients, in which a

20

25

probe labeled with any suitable tracer, and comprising all or part of SEQ ID NO. 3, is used.

- 14. Recombinant protein obtained using an expression cassette in a bacterial host, characterized in that its protein sequence consists of SEQ ID NO. 31.
- 15. Protein according to claim 14, characterized in that the bacterial host is *E. coli*.
 - 16. Reagent for detecting, in a biological sample, an autoimmune disease or monitoring pregnancy, comprising at least one fragment according to any of claims 1 one to 6, transcription/translation product, as obtained according to the method of claim 11, one synthetic peptide belonging to SEQ ID NO. 31 or one protein according to claim 15 or 16.
 - 17. Use of a fragment according to any one of claims 1 to 6, of a transcription/translation product, as obtained according to the method of claim 11, of a synthetic peptide belonging to SEQ ID NO. 31 or of a protein according to claim 15 or 16, for detecting, in a biological sample, susceptibility to an autoimmune disease, or monitoring pregnancy.
- 18. Use according to claim 17, characterized in that the autoimmune disease is multiple sclerosis.

FIG 1

1	ORFs	538	538))	OH P	413 and 305
Ċ	repears	yes	yes	yes	yes	Plo Du
eimilarition			%96	88%	%69%	88%
N Supplemental	Boogs DAIA	Vecoils RINA	PG083M05 [7]	BAC378 [14]	Q11M15 [21]	U134E6 [x]
	7582		37879	14079	27999	94627
			į			
₹]	· 					i
_T.		28274	6911	35199	21299	

REPLACEMENT SHEET (RULE 26)

1 / 5

FIG 1

7582
97879
— 14079 BA
27999
94627

REPLACEMENT SHEET (RULE 26)

PCT/FR00/00144

FIG 2

10 20 30 40 50 1234567890 1234567890 1234567890 1234567890	
OCTAGAAGT ATTCTGGAGA ATTGGGACCA ATGTGACACT CAGACGCTAA PRTYSGELGPM HSDA K LEPILEN W D Q C D T Q T L R . N V F W R I G T N V T L R R	
CAANGAAACG ATTTATATTIC TICTOCAGTA COOCCTGGCC ACAATATCCT KETIYILLQYRLA TISS KKRFIFFCSTAWPQYP ERNDLYSSAVPPGHNIL	100
CTTCAAGGGA GAGAAACCTG OCTTCCTGAG GGAAGTATAA ATTATAACAT SRERNLAS.GKYKL.H LQGRETWLPEGSINYNI FKGEKPGFLREV.IITS	150
CATCITACAG CTAGACCICT TCTGTAGAAA GGAGGGCAAA TGGAGTGAAG H L T A R P L L . K G G Q M E . S I L Q L D L F C R K E G K W S E V S Y S . T S S V E R R A N G V K	200
TOOCATATOT OCAAACTTIC TTITCATTAA CAGACAACTC ACAATTATOT A I C A N F L F I K R Q L T I M . P Y V Q T F F S L R D N S Q L C C H M C K L S F H . E T T H N Y V	250
AAAAAGIGIG GITTATOCCC TACAGGAACC CCTCAGAGTC CACCTCCCTA K V W F M P Y R K P S E S T S L K K C G L C P T G S P Q S P P P Y K S V V Y A L Q E A L R V H L P T	300
COCCACCIC COCTOCCCA CICCITCIC ANCIATAG CACCCCTT PQRPLPDSFLNGPPF PSVPSPTPSSTNKDPPL PASPPRLLPQLIRTPL	350
TAAOOCAAAC GCTICCAAAAG GACATACACA AAGGGTAAA CAATGAACCA N P N G P K G D R Q R G K Q . T K T Q T V Q K E I D K G V N N E P . P K R S K R R . T K G . T M N Q	400
AACAGIOOCA ATATTOOOCG ATTATOOCCC CTOCAAOCAG TGACAGGAGG ECQYSPIMPPPSSERR KSANIPRLCPLQAVRGG RVPIFPDYAPSKQ.EEE	45 0
VEAATTOOOC CO2000200G TOOCTGTACC TTTTTCTCTC TCAGACTTAA R I R P S Q S A C T F F S L R L K E F G P A R V P V P F S L S D L K N S A Q P E C L Y L F L S Q T .	500

REPLACEMENT SHEET (RULE 26)

FIG 2 (continued)

10 20 30 40 50 1234567890 1234567890 1234567890 1234567890	
ACCAAATTAA AATAGACCTA OOTAAATTCT CAGATAACCC TGACOOCTAT A N N R P R . I L R P R L Y Q I K I D L G K P S D N P D G Y S K L K T V N S Q I T L T A I	550
ATTIGATIGHT TACAACCIT ACGACAATCC THTICATICTGA CATGGAGAGA . C F T R V R T I L . S D M E R I D V L Q G L G Q S F D L T W R D L M F Y K G . D N P L I . H G E I	600
TATAATGITA CTACTAAATC AGACACTAAC CCCAAATGAG AGAAGIGOCG Y N V T T K S D T N P K . E K C R I M L L L N O T L T P N E R S A A . C Y Y . I R H . P Q M R E V P	650
CTGTAACTOC AGCCCGAGAG TITGCCGATC TITGGTATCT CAGTCAGGCCCCAGAG TITGCCGATC TITGGTATCT CAGTCAGGCCCCCCCCCCCCCCCCCCCCCCC	700
AACAATAGGA TGACAACAGA GGAAAGAACA ACTCCCACAG GCCAGCAGGC Q . D D N R G K N N S H R P A G N N R M T T E E R T T P T G Q Q A T I G . Q Q R K E Q L P Q A S R Q	750
AGTTOCCAGT GIAGACCCIC ATTOCCACAC AGAATCAGAA CATGCAGATT SSQCRPSLGHRIRT WRL VPSVDPHWDTESEHGDW FPV.TLIGTQNQNMEI	800
GSTGCCACAA ACATTIGCIA ACTIGOGIGC TAGAAGGACT GAGGAAAACT V P Q T F A N L R A R R T E E N . C H K H L L T C V L E G L R K T G A T N I C . L A C . K D . G K L	850
AGGACAACC CTATGAATTA CICAATGAIG TOCACTATAA CACAGGGAAA EEAYELLNDVHYNTGK RKKPMNYSMMSTITQGK GRSL.ITQ.CPL.HRER	900
GEANGAAAT CITACIGCTT TICTGCACG ACTAACGCAG GCATTGCACGA G R K S Y C F S G Q T K G G I E E E E N L T A F L D R L R E A L R K K K I L L L F W T D . G R H . G	950
ACCATACCIC CCIGICACCI CACICIAITG AACCCAACT AATCTTAAAG A Y L P V T . L Y . R P T N L K G H T S L S P D S I E G Q L I L K S I P P C H L T L L K A N . S . R	1000

REPLACEMENT SHEET (RULE 26)

FIG 2 (continued)

10 20 30 40 50 1234567890 1234567890 1234567890 1234567890	
CATAAGTTA TCACTCAGIC ACCTCACAC ATTAGAAAAA ACTTCAAAAG . V Y H S V S C R H . K K L Q K D K F I T Q S A A D I R K N F K S I S L S L S Q L Q T L E K T S K V	1050
TOTOCCTTAG OCCOGGASCA GAACTIAGAA ACCCIATTTA ACTTGSCATC S A L G P E Q N L E T L F N I, A S L P . A R S R T . K P Y L T W H P C L R P G A E L R N P I . L G I	1100
CTCAGTTTTT TATAATAGAG ATCAGGAGGA GCAGGGGAAA CGGGACAAAC S V F Y N R D Q E E Q A K R D K R Q F F I I E I R R S R R N G T N L S F L R S G G A G E T G Q T	1150
COCCATAAAAA AAAAAACCCCC COCTCACCCCCCCCCC	1200
AAGCACACIT TOGAGOCTCT GCAAAAGGGA AAAGCTGGGC AAATCAAATG KQTLEALQKG KAGQIKC SRLWRLC KREKLG KSNA ADF GGS AKGKSWA NQM	1250
CCTAATAGOG CTOCCTTCCA GTCCCGTCTA CAAGGACACT TTAAAAAAGA LIG LASSAVY KDT LKKIG W LP V R S T R T L . K R P N R A G F Q C G L Q G H F K K D	1300
TIATOCAAGI AGAAATAAGC OGCOOCCIIG TOCAIGOOCC TIAGGICAAG I Q V E I S R P L V H A P Y V K L S K . K . A A P L S M P L T S R Y P S R N K P P P C P C P L R Q G	1350
GCAATCACTG CAACGCCCAC TOCCCCACGC CATCAACATA CTCTCACTCA G I T G R P T A P G D E D T L S Q E S L E G P L P Q G M K I L . V R N H W K A H C P R G . R Y S E S	1400
CAMOCCATTA ACCAGATGAT COMOCACCAG GACTGAGGGT GOCCGGGGGG : K P L T R . S S S R T E G A R G E S H . P D D P A A G L R V P G A E A I N Q M I Q Q Q D . G C P G R	1450
ASOSCCACC CATGODATCA COCTCACACA GOODOGGTA TGITTGACCA R Q P M P S P S Q S P G Y V . P S A S P C H H P H R A P G M F D H A P A H A I T L T E P R V C L T I	500

REPLACEMENT SHEET (RULE 26)

PCT/FR00/00144

FIG 2 (continued)

10 20 30 40 50

1234567890 1234567890 1234567890 1234567890

TICACACCCA A 1511

L R A

. E P

F. S Q

DECLARATION AND POWER OF ATTORNEY UNDER 35 USC §371(c)(4) FOR PCT APPLICATION FOR UNITED STATES PATENT

As a below named inventor, I hereby declare that: my residence, post office address and citizenship are as stated below under my name;

I verily believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought, namely the invention entitled: Endogeneous nucleic acid fragment associated with an autoimmune disease, labeling method and reagent

described and claimed in international application number PCT FR00/00144 filed 21th January 2000.

I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations §1.56.

Under Title 35, U.S. Code §119, the priority benefits of the following foreign application(s) filed by me or my legal representatives or assigns within one year prior to my international application are hereby claimed:

France Nº 99/00888 filed on 21th January 1999

The following application(s) for patent or inventor's certificate on this invention were filed in countries foreign to the United States of America either (a) more than one year prior to my international application, or (b) before the filing date of the above-named foreign priority application(s):

I hereby appoint the following as my attorneys of record with full power of substitution and revocation to prosecute this application and to transact all business in the Patent Office:

James A. Oliff, Reg. No. 27,075; William P. Berridge, Reg. No. 30,024; Kirk M. Hudson, Reg. No. 27,562; Thomas J. Pardini, Reg. No. 30,411; Edward P. Walker, Reg. No. 31,450; Robert A. Miller, Reg. No. 32,771; Mario A. Costantino, Reg. No. 33,565; Caroline D. Dennison, Reg. No. 34,494; and Stephen J. Roe, Reg. No. 34,463.

ALL CORRESPONDENCE IN CONNECTION WITH THIS APPLICATION SHOULD BE SENT TO OLIFF & BERRIDGE, PLC, P.O. BOX 19928, ALEXANDRIA, VIRGINIA 22320, TELEPHONE (703) 836-6400.

I hereby declare that I have reviewed and understand the contents of this Declaration, and that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

1	Typewritten F of Sole or Firs	ull Name /{	90			
	of Sole or Firs	st Inventor /	Glaucia		PARANHOS-BACCALA	
2	Inventor's Si		Given Name GLAUCIA	Middle Initial PARAWHOS	Family Name BACCALA	
3	Date of Signa	ature _	<i>o</i> 7	18	2001	
	_	_	Month	Day	Year	
	Residence:	6900	3 LYON		France	
	City Citizenship: SWISS		City	State or Province	Country	
Post Office Address:			75 Cours Gam	betta 69003 LYON (Fr	ance)	
	•	ert complete mailing ess, including country)			

Note to Inventor: Please sign name on line 2 exactly as it appears in line 1 and insert the actual date of signing on line 3.

(Discard this page in a sole inventor application)

		0-00		
1	Typewritten Full Name	<i>Y</i> .		
	of Joint Inventor	François • Given Name		MALLET
2	Inventor's Signature:	frum (01)	Middle Initial	Family Name
3	Date of Signature:		17	2001
	Residence: 69100 V	Month ILLEURBANNE	Day	Year France
	Citizenship: French	City	State or Province	Country
	Post Office Address: (Insert complete mailing address, including country)		nce 69100 VILLEURBANNE	(France)
1	Typewritten Full Name of Joint Inventor	Cécile	<u> </u>	VOISSET
2	Inventor's Signature:	Given Name	Middle Initial	Family Name
3	Date of Signature:	Ø 8	Ol	2001
] D	Residence: LONI	Month DON WC1H 9QW	h Day	Year GB
i M	Citizenship: French	City	State or Province	Country
	Post Office Address:	80 Clare Court Judd	Street – LONDON WC1H 9Q	W (GR)
	(Insert complete mailing address, including country)		John Boll Welling	, (G <i>L</i>)
_11	Typewritten Full Name of Joint Inventor			
2 -3 -1	Inventor's Signature:	Given Name	Middle Initial	Family Name
3	Date of Signature:			
	Residence:	Month	Day	Year
10000 110000 110000	Citizenship:	City	State or Province	Country
	Post Office Address: (Insert complete mailing address, including country)			
1	Typewritten Full Name of Joint Inventor			
2	Inventor's Signature:	Given Name	Middle Initial	Family Name
3	Date of Signature:			
	Residence:	Month	Day	Year
	Citizenship:	City	State or Province	Country
	Post Office Address: (Insert complete mailing address, including country)	-		

Note to Inventor: Please sign name on line 2 exactly as it appears in line 1 and insert the actual date of signing on line 3.

This form may be executed only when attached to the first page of the Declaration and Power of Attorney of the application to which it pertains.

<213> Homo sapiens

```
SEQUENCE LISTING
<110> BIO MERIEUX
<120> Endogenous nucleic acid fragment associated with an
      autoimmune disease, labeling method and reagent
<130> SEP19-endogenous
<140>
<141>
<150> FR 9900888
<151> 1999-01-21
<160> 31
<170> PatentIn Ver. 2.1
<210> 1
<211> 1511
<212> DNA
<213> Homo sapiens
cctagaacgt attctggaga attgggacca atgtgacact cagacgctaa gaaagaaacg 60
atttatattc ttctgcagta ccgcctggcc acaatatcct cttcaaggga gagaaacctg 120
getteetgag ggaagtataa attataacat catettacag etagacetet tetgtagaaa 180
ggagggcaaa tggagtgaag tgccatatgt gcaaactttc ttttcattaa gagacaactc 240
acaattatgt aaaaagtgtg gtttatgccc tacaggaagc cctcagagtc cacctcccta 300
ceccagegte ceetecega etectteete aactaataag gaceeceett taacccaaac 360
ggtccaaaag gagatagaca aaggggtaaa caatgaacca aagagtgcca atattccccg 420
attatgecee etceaageag tgagaggagg agaattegge ecagecagag tgeetgtace 480
tttttctctc tcagacttaa agcaaattaa aatagaccta ggtaaattct cagataaccc 540
tgacggctat attgatgttt tacaagggtt aggacaatcc tttgatctga catggagaga 600
tataatgtta ctactaaatc agacactaac cccaaatgag agaagtgccg ctgtaactgc 660
agcecgagag tttggcgate tttggtatet eagteaggee aacaatagga tgacaacaga 720
ggaaagaaca acteccacag gecageagge agtteccagt gtagaccete attgggacae 780
agaatcagaa catggagatt ggtgccacaa acatttgcta acttgcgtgc tagaaggact 840
gaggaaaact aggaagaagc ctatgaatta ctcaatgatg tccactataa cacagggaaa 900
ggaagaaaat cttactgctt ttctggacag actaagggag gcattgagga agcatacctc 960
cetgteacet gaetetattg aaggeeaact aatettaaag gataagttta teacteagte 1020
agetgeagae attagaaaaa aetteaaaag tetgeettag geeeggagea gaaettagaa 1080
accetattta acttggcate etcagttttt tataatagag atcaggagga geaggegaaa 1140
egggacaaac gggataaaaa aaaaaggggg ggtccactac tttagtcatg gccctcaggc 1200
aagcagactt tggaggetet gcaaaaggga aaagctggge aaatcaaatg cetaataggg 1260
ctggetteca gtgcggteta caaggacact ttaaaaaaga ttatccaagt agaaataagc 1320
egececettg tecatgecee ttacgtcaag ggaatcactg gaaggeecae tgececaggg 1380
gatgaagata ctctgagtca gaagccatta accagatgat ccagcagcag gactgagggt 1440
georgggeg agegerage catgeratea certearaga georgggta tgtttgacca 1500
ttgagagcca a
<210> 2
<211> 2009
<212> DNA
```

```
atacgactae tatagggega attgggeeet etagatgeat getegagegg eegeeagtgt 60
gatggatate tgcagaatte geeetttgte egetgtgete etgateeage gaggegeeea 120
ttgctgctcc caattgggct aaaggcttgc cattgttccc acacggctaa gtgcccgggt 180
tcatcctaat tgagctgaac actagtcact gggttccatg gttctcttcc atgacccacg 240
gottotaata gagototaat actoaccaca tggoocaaga ttocattoot tggaatoogt 300
gaggccaaga accccaggtc agagaacacg aggcttgcca ccgtcttgga agtggcccqc 360
egecatettg ggagetetgg gageaaggae eecceagtaa cattttggea accacaaagq 420
gacetecaaa gegatgggaa acatteeece caaggeaaaa aegeeectaa gatqtattet 480
qqaqaattgg gaccaatgtg acactcagac gctaagaaag aaacgattta tattcttctg 540
cagtacogco tggccacaat atcotottoa agggagagaa acctggctto otgagggaag 600
tataaattat aacatcatct tacagetaga cetettetgt agaaaggagg geaaatggag 660
tqaaqtqcca tatgtgcaaa ctttcttttc attaagagac aactcacaat tatgtaaaaa 720
gtgtggttta tgccctacag gaagccctca gagtccacct ccctacccca gcgtcccccc 780
congacted testeaacta ataaggaces contituade caaanggtes aaaaggagat 840
agacaaaggg gtaaacaatg aaccaaagag tgccaatatt ccccgattat gccccctcca 900
agcagtgaga ggaggagaat teggeecage cagagtgeet gtacettttt eteteteaga 960
cttaaagcaa attaaaatag acctaggtaa attctcagat aaccctgacg gctatattga 1020
tgttttacaa gggttaggac aatootttga totgacatgg agagatataa tgttactact 1080
aaatcagaca ctaaccccaa atgagagaag tgccgctgta actgcagccc gagagtttgg 1140
egatetttgg tateteagte aggteaacaa taggatgaca acagaggaaa gaacaactee 1200
cacaqqccaq caggcagttc ccagtgtaga ccctcattgg gacacagaat cagaacatgg 1260
agattggtgc cacaaacatt tgctaacttg cgtgctagaa ggactgagga aaactaggaa 1320
gaagcctatg aattactcaa tgatgtccac tataacacag ggaaaggaag aaaatcctac 1380
tgcttttctg gacagactaa gggaggcatt gaggaagcat acctccctgt cacctgactc 1440
tattgaagge caactaatct taaaggataa gtttatcact cagtcagctg cagacattag 1500
aaaaaacttc aaaagtctgc cttaggcccg gagcagaact tagaaaccct atttaacttg 1560
gcatcotcag tittitataa tagagatcag gaggagcagg cgaaacggga caaacgggat 1620
aaaaaaaaa gggggggtcc actactttag tcatggccct caggcaagca gactttggag 1680
getictggaaa agggaaaagc tgggcaaatc aaatgcctaa tagggctggc ttccagtgcg 1740
gtctacaagg acactttaaa aaagattatc caagtagaaa taagccgccc ccttgtccat 1800
geocettacg teaagggaat cactggaagg cecactgeee caggggatga agatactetg 1860
agtcagaage cattaaccag atgatecage ageaggactg agggtgeeeg gggegagege 1920
cageccatge cateaceetn acagageece gggtatgett gaccattgag agccaggagg 1980
                                                                  2009
ttaactgtct cctggacact ggcgcagcc
<210> 3
<211> 1056
<212> DNA
<213> Homo sapiens
<400> 3
ctagaacgta ttctggagaa ttgggaccaa tgtgacactc agacgctaag aaagaaacga 60
tttatattet tetgeagtae egeetggeea caatateete tteaagggag agaaacetgg 120
cttcctgagg gaagtataaa ttataacatc atcttacagc tagacctctt ctgtagaaag 180
gagggcaaat ggagtgaagt gccatatgtg caaactttct tttcattaag agacaactca 240
caattatgta aaaagtgtgg tttatgccct acaggaagec etcagagtee acetecctae 300
cocagogico cotococgao toottootoa actaataagg accoccotti aaccoaaacg 360
gtccaaaagg agatagacaa aggggtaaac aatgaaccaa agagtgccaa tattccccga 420
ttatgeecee tecaagcagt gagaggagga gaatteggee cagecagagt geetgtaeet 480
ttttctctct cagacttaaa gcaaattaaa atagacctag gtaaattctc agataaccct 540
gacggctata ttgatgtttt acaagggtta ggacaatcct ttgatctgac atggagagat 600
ataatgttac tactaaatca gacactaacc ccaaatgaga gaagtgccgc tgtaactgca 660
geocgagagt ttggcgatet ttggtatete aqteaggeca acaataggat gacaacagag 720
gaaagaacaa ctcccacagg ccagcaggca qttcccagtg tagaccctca ttgggacaca 780
gaatcagaac atggagattg gtgccacaaa catttgctaa cttgcgtgct agaaggactg 840
```

aggaaaacta ggaagaagcc tatgaattac teaatgatgt ccactataac acagggaaag 900

ctgtcacctg	ttactgettt actetattga ttagaaaaaa	aggccaacta	ctaagggagg atcttaaagg ctgcct	cattgaggaa ataagtttat	gcatacctcc cactcagtca	960 1020 1056
<210> 4 <211> 25 <212> DNA ·213> Homo	sapiens					
<400> 4 cggacatcca	aagtgatggg	aaacg				25
<210> 5 <211> 26 <212> DNA						
<213> Homo	sapiens					
<400> 5 ggacaggaaa	gtaagactga g	gaaggc			:	26
<210> 6 <211> 26						
<211> 20						
<213> Homo	sapiens					
<400> 6						
cctagaacqt a	attctggaga a	ittggg			2	6
<210> 7 <211> 26 <212> DNA						
<213> Homo s	sapiens					
<400> 7 tggctctcaa t	ggtcaaaca t	acccg			2	6
<210> 8 <211> 21 <212> DNA						
<213> Homo s	apiens					
<400> 8 atgggaaacg t	teccceega g				2:	1
<210> 9						
<211> 21 <212> DNA						
<213> Homo s	apiens					
<400> 9						
gcctaaggc ad	gacttttga a				2.2	

<210> 10 <211> 409 <212> PRT <213> Homo sapiens

<400> 10

Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro 1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Ile Met Gly Asn Ile Pro Pro Lys Ala Lys Thr Pro Leu Arg Cys Ile
35 40 45

Leu Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu Arg Lys Lys Arg 50 55 60

Phe Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr Pro Leu Gln Gly 65 70 75 80

Arg Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr Asn Ile Ile Leu 85 90 95

Gln Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp Ser Glu Val Pro 100 105 110

Tyr Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser Gln Leu Cys Lys 115 120 125

Lys Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser Pro Pro Pro Tyr 130 135 140

Pro Ser Val Pro Pro Pro Thr Pro Ser Ser Thr Asn Lys Asp Pro Pro 145 150 155

Leu Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly Val Asn Asn Glu 165 170 175

Pro Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro Leu Gln Ala Val Arg 180 185 190

Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro Phe Ser Leu Ser 195 200 205

Asp Leu Lys Gln lle Lys Ile Asp Leu Gly Lys Phe Ser Asp Asn Pro 210 215 220

Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln Ser Phe Asp Leu 225 230 235 240

Thr Trp Arg Asp Ile Met Leu Leu Leu Asn Gln Thr Leu Thr Pro Asn 245 250 255

Glu Arg Ser Ala Ala Val Thr Ala Ala Arg Glu Phe Gly Asp Leu Trp 260 265 270 Tyr Leu Ser Gln Val Asn Asn Arg Met Thr Thr Glu Glu Arg Thr Thr 280

Pro Thr Gly Gln Gln Ala Val Pro Ser Val Asp Pro His Trp Asp Thr

Glu Ser Glu His Gly Asp Trp Cys His Lys His Leu Leu Thr Cys Val

Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Pro Met Asn Tyr Ser Met

Met Ser Thr Ile Thr Gln Gly Lys Glu Glu Asn Pro Thr Ala Phe Leu 345

Asp Arg Leu Arg Glu Ala Leu Arg Lys His Thr Ser Leu Ser Pro Asp 360

Ser Ile Glu Gly Gln Leu Ile Leu Lys Asp Lys Phe Ile Thr Gln Ser 375

Ala Ala Asp Ile Arg Lys Asn Phe Lys Ser Leu Pro Lys Leu Ala Ala 395

Ala Leu Glu His His His His His 405

<210> 11

<211> 393

<212> PRT

<213> Homo sapiens

<400> 11

Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Ile Met Gly Asn

Ile Pro Pro Lys Ala Lys Thr Pro Leu Arg Cys Ile Leu Glu Arg Ile 2.0 25

Leu Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu Arg Lys Lys Arg 40

Phe Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr Pro Leu Gln Gly 50

Arg Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr Asn Ile Ile Leu

Gin Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp Ser Glu Val Pro

Tyr Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser Gln Leu Cys Lys

Lys Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser Pro Pro Pro Tyr 120

Pro Ser Val Pro Ser Pro Thr Pro Ser Ser Thr Asn Lys Asp Pro Pro

Leu Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly Val Asn Asn Glu 150 155

Pro Lys Ser Ala Asn lle Pro Arg Leu Cys Pro Leu Gln Ala Val Arg

Gly Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro Phe Ser Leu Ser 185

Asp Leu Lys Gln lle Lys Ile Asp Leu Gly Lys Phe Ser Asp Asn Pro 200

Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln Ser Phe Asp Leu 215

Thr Trp Arg Asp Ile Met Leu Leu Leu Asn Gln Thr Leu Thr Pro Asn 235

Glu Arg Ser Ala Ala Val Thr Ala Ala Arg Glu Phe Gly Asp Leu Trp 250

Tyr Leu Ser Gln Ala Asn Asn Arg Met Thr Thr Glu Glu Arg Thr Thr 265

Pro Thr Gly Gln Gln Ala Val Pro Ser Val Asp Pro His Trp Asp Thr 280

Glu Ser Glu His Gly Asp Trp Cys His Lys His Leu Leu Thr Cys Val 295

Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Pro Met Asn Tyr Ser Met 310

Met Ser Thr Ile Thr Gln Gly Lys Glu Glu Asn Leu Thr Ala Phe Leu 325

Asp Arg Leu Arg Glu Ala Leu Arg Lys His Thr Ser Leu Ser Pro Asp 345

Ser Ile Glu Gly Gln Leu lle Leu Lys Asp Lys Phe Ile Thr Gln Ser 360

Ala Ala Asp Ile Arg Lys Asn Phe Lys Ser Leu Pro Lys Leu Ala Ala 375

Ala Leu Glu His His His His His 390

<210> 12

<211> 26

<212> DNA

<213> Homo sapiens

<400> 12 ctagaacgta ttctggagaa ttggga	26
<210	
<400> 13 cctaaggcag acttttgaag	20
<210> 14 <211> 54 <212> DNA	
<213> Homo sapiens	
<400> 14 tttggtaata cgactcacta tagggcagcc accatgggaa acgttccccc cgag	54
<210> 15 <211> 44 <212> DNA	
<213> Homo sapiens	
<400> 15 tttttttttttttttc aggetgegee agtgtecagg agae	44
<210> 16 <211> 20 <212> DNA	
<213> Homo sapiens	
<400> 16 tgtccgctgt gctcctgatc	20
<210> 17 <211> 44 <212> DNA	
<213> Homo sapiens	
<400> 17 ttttttttt tttttttc aggetgegee agtgteeagg agae	44
<210> 18 <211> 678 <212> DNA	
<213> Homo sapiens	
<400> 18 tcagggatag cocceatota titggocagg cartagecca agactigage cagticicat acciggatat tottgiocht iggrafgeg argatifact titaggogge catteagas	60

```
Cottgtgcca tcaagccacc caagtgctct taaatttcct cgccacctgt ggctacaagg 180
tttccaaacc aaaggctcag ctctgctcac agcagaaggc tatttaccct aaatacttag 240
ggctgaaatt atccaaaggc accagggccc tcagtgagga atgtatccag cctatactgg 300
cttatcctta tcccaaaacc ctaaaacaac taagaaggtt ccttggcata ataggcataa 360
caggcataac aggtttctgc tgaatatgga ttcccaagta cggcaaaata gccagaccat 420
tatatacact aattaaggaa actcagaaag ccaataccca tttagtaaga tggacacctg 480
aagcagaggo agotttocag googtaaaga acaccotaac ocaagcocca gtgttaaqot 540
tgccageggg gcaagacttt tctttctgtg tcacagaaaa aataggaata gctntaggaq 600
teettacaca ggteegaggg accagettge aacceatgge atacetgagt aaggaaattg 660
atgtagtggc aaagggtt
<210> 19
<211> 591
<212> DNA
<213> Homo sapiens
<400> 19
ccatggccat ctacactgaa caagatttat acaatcatgt cgtacctaag ccccacaaca 60
aaagagtace cattetteet titgitatea gageaggagt getaggeaga etaggtaetg 120
gcattggcag tatcacaacc tctactcagt tctactacaa actatctcaa gaaataaatg 180
gtgacatgga acaggtcact gactccctgg tcaccttgca agatcaactt aactccctag 240
cagcagtagt cetteaaaat egaagagett tagaettget aacegeeaaa agagggggaa 300
cotgettatt tetaggagaa gaacgorger attargetaa toaatooaga attgecactg 360
agaaagttaa agaaattcga gatcgaatac aatgtagagc agaggagctt caaaacaccg 420
aacgetgggg cetectcage caatggatge cetgggttet cecettetta ggacetetag 480
cagetetaat attgttacte etetttggae eetgtatett taaceteett gttaagtttg 540
totottocag aattgaaget gtaaagetac agatggtott acaaatctag a
<210> 20
<211> 1321
<212> DNA
<213> Homo sapiens
<400> 20
caacaatcgg gatataaacc caggcattcg agctggcaac agcagccccc ctttgggtcc 60
cttccctttg tatgggagct gttttcatgc tatttcactc tattaaatct tgcaactgca 120
ctcttctggt ccatgtttct tacggctcga gctgagcttt tgctcaccgt ccaccactgc 180
tgtttqccac caccqcaqac ctgccqctga ctcccatccc tctggatcct gcagggtgtc 240
cgctgtgctc ctgatccagc gaagcgccca ttgccgctcc caattgggct aaaggcttgc 300
cattgttcct gcacggctaa gtgcctgggt ttgttctaat tgagctgaac actagtcact 360
gggttccatg gttctcttct gtgacccacg gcttctaata gaactataac acttaccaca 420
tggcccaaga ttccattcct tggaatccgt gaggccaaga actccaggtc agagaatacg 480
aagettgeca ceatettgga ageggeetge taccatettg gaagtggtte accaecatet 540
tgggagctct gtgagcaagg acccccggt aacattttgg caaccacgaa cggacatcca 600
aagtgatggg aaacgtteee egeaagacaa aaacgeeeet aagaegtatt etggaaaatt 660
gggaacaatt tgaccetcag acactaagaa agaaacgact tatattette tgcagtgeeg 720
cctggcactc ctgagggaag tataaattat aacaccatct tacagctaga cctcttttgt 780
agaaaaggca aatggagtga agtgccataa gtacaaactt tcttttcatt aagagacaac 840
tcacaattat gtaaaaagtg tgatttatgc cctacaggaa gccttcagag tctacctccc 900
tatoccagca tocccgacto ottocccact taataaggac coccettoaa cocaaatggt 960
ccaaaaggag atagacaaaa gggtaaacag tgaaccaaag agtgccaata ttccccaatt 1020
atgacccctc caagcagtgg gaggaagaga attcggccca gccagagtgc atgtgccttt 1080
ttctctccca qacttaaagc aaataaaaac agacttaggt aaattctcag ataaccctga 1140
tggctatatt ggtgttttac aagggttagg acaattettt gatetgacat ggagagatat 1200
atatgtcact gctaaatcag acactaaccc caaatgagag aagtgccacc ataactgcag 1260
cctgagagtt tggcgatctc tggtatctca gtcaggtcaa tgataggatg acaacagagg 1320
```

a

1321

<210> 21 <211> 2938 <212> DNA <213> Homo sapiens <400> 21 caacgacgga catccaaagt gatgggaaac gttccccgca agacaaaaac gcccctaaga 60 cgtattctgg agaattggga ccaatttgac cctcagacac taagaaagaa acgacttata 120 ttcttctgca gtgccgcctg gcactcctga gggaagtata aattataaca ccatcttaca 180 gctagactic tittgtagaa aaggcaaatg gagtgaagtg ccataagtac aaacttictt 240 ttcattaaga gacaactcac aattatgtaa aaagtgtgat ttatgcccta caggaagcct 300 tragaginta colocitate cragnation egactrotte recaactaat aaggaccere 360 cttcaaccca aatggtccaa aaggagatag acaaaagggt aaacagtgaa ccaaagagtg 420 ccaatattcc ccaattatga ccctcccaa gcagtgggag gaagagattc ggcccagcca 480 gagtgcatgt gctttttctt ctcccagact taaagcaaat aaaaacagac ftaggtaaat 540 totcagataa tootgatggo tatattgatg ttttacaagg gttaggacaa ttotttgato 600 tgacatggag agatataatg tcactgctaa atcagacact aaccccaaat gagagaagtg 660 ccaccataac tgcagcctga gagtttggcg atctctggta tctcagtcag gtcaatgata 720 ggatgacaac agaggaaaga gatgatcccc acagccagca agcagttccc agtctasacc 780 ctcattgggg acacagaaat cagtaacatg ggagattggt gctgcagaca tttgctaact 840 tgtgtgctac aaggactaag gaaaactacg aagaaaatct acgaattact caatgatgtc 900 caccataaca caggggaagg gaagaaaatc ctactgcctt tctggagaga ctaagggagg 960 cattgaggaa gcgtgcctct ctgtcacctg actcttctga aggccaacta atcttaaagc 1020 gtaagtttat cactcagtca gctgcagaca ttagaaaaaa cttcaaaagt ctgccqtaqq 1080 cccggagcaa aacttagaaa ccctattgaa cttggcaacy tcggtttttt ataatagaga 1140 traggaggag raggreggaar aggaraaarg ggattaaaaa aaaggreare getttagtra 1200 tgaccetcag geaagtggae tttggagget etggaaaagg gaaaagetgg geaaattgaa 1260 tgcctaatag ggcttgcttc cagtgcggtc tacaaggaca ctttaaaaaa gattgtccaa 1320 gtagaagtaa gccgcccctt cgtccatgcc ccttatttca agggaatcac tggaaggccc 1380 actgccccag gggacaaagg tcttttgagt cagaagccac taaccagatg atccagcagc 1440 aggactgagg gtgcctgggg caagegccat cecatgccat cacceteaca gagecetggg 1500 tatgcttgac cattgagggc caggaaggtt gtctcctgga cactggtgcg gtcttcttag 1560 tottactott otgtocogga caactgtoot ocagatotgt cactatotga gggggtoota 1620 agacgggcag tcactagata cttctcccag ccactaagtt atgactgggg agctttattc 1680 ttttcacatg cttttctaat tatgcttgaa agccccacta ccttgttagg gagagacatt 1740 ctagcaaaag caggggccat tatacacctg aacataggag aaggaacacc cgtttgttgt 1800 eccetgettg aggaaggaat taateetgaa gtetgggeaa cagaaggaea atatggaega 1860 gcaaagaatg coogtootgt toaagttaaa ctaaaggatt coacttoott tooctaccaa 1920 aggcagtace coetcagace caaggcecaa caaggattee aaaagattgt taaggactta 1980 aaagcccaag gcttagtaaa accatgcata actccctgca gtaattccgt agtggattga 2040 ggaggcacag aaacccagtg gacagtggag ggttagtgca agatctcagg attatcaatg 2100 gaggccgttg tccttttata cccagctgta cctagccctt atactgtgct ttcccaaata 2160 ccagaggaag cagagtggtt tacacteetg gacettaagg atgeettett etgeateeet 2220 gracatectg acteteaatt ettgettegee tetgaagata etteaaacce aacateteaa 2280 ctcacctgga ctgttttacc ccaagggttc agggatagcc cccatctatt tggccaggca 2340 ttageceaag aettgageca ateetratae etggacaett gteetteggt aggtggatga 2400 tttacttttg gccgcccatt cagaaacctt gtgccatcaa gccacccaag cgctcttcaa 2460 tttecteget acetgtgget acatggttte caaaccaaag geteaactet geteacagca 2520 ggttacttag ggctaaaatt atccaaaggc accagggccc tcagtgagga acacatccag 2580 cctatactgg cttatcctca tcccaaaacc ctaaagcaac taaggggatt ccttggcgta 2640 ataggtttct gccgaaaatg gattcccagg tttggcgaaa tagccaggtc attaaataca 2700 ctaattaagg aaactcagaa agccaatacc catttagtaa gatggacaac tgaagtagaa 2760 gtggctttcc aggccctaac ccaagcccca gtgttaagtt tgccaacagg gcaagacttt 2820 tetteatatg teacagaaaa aacaggaata getetaggag teettacaca gateegaggg 2880 atgagettge aacetgtgge gtacetgaet aaggaaattg atgtagtgge aaagggtt

```
<210> 22
 <211> 1422
 <212> DNA
 <213> Homo sapiens
 <400> 22
 tcagggatag cccccatcta tttggccagg cattagccca agacttgagt cagttatcat 60
 acctggacae tettgteett cagtatgtgg atgatttaet tttagetgee tgttcagaaa 120
 cettgtgeca teaagecace caageactet taaattteet egecacetgt ggetacaagg 180
 tttccaaaga gaageteage tetgeteaca geaggttaaa taettaggae taagattate 240
 caaaggcacc aaggccctca gtgaggaatg tatccagcct atactggctt atcctcatct 300
 caaaacccta aagcaactaa gagagtteet tggcataaca ggettetgee gaatatggat 360
 tccccaggta tggcaaaata gccaggccat tatatacagt aattaaggaa actcagaaag 420
 ccaataccca tttaataaga tggatacctg aagccaaagt ggctttccag gcccctaaag 480
 aaggeettaa acceaagtee eagtgttaag ettgeeaacg gggeaagaet tttetttata 540
 catcacagaa aaaaacagaa acagctctgg gagtccttac acaggtccaa gggacgagct 600
tgcaacccat ggcatacctg agtaaggaaa ctgatgtagt ggcaaagggt tggcttcatt 660
gtttatgggt agtggtggca gtagcagttg tagtatctga agcagttaaa ataatacagg 720
ggagagatet tactgtgtgg acateteatg aggtgaacag catacteact getaaaggag 780
acttgtggct gtcagacaac cgtttactta aatatcaggc tctattactt gaaaggccag 840
tgctgcaact gtgcacttgt gcaactctta acccagtcnc atttcttcca gacaatgaag 900
atagaatata actgtcaaca aataatttct caaacctatg ccactcgagg ggaccttcta 960
gaagtteect tgaetgatee tgaeetteaa ettgtataet gatggaagtt eetttgtaga 1020
aaaaggactt caaaagcggg gtatgcagtg gtcagtgata atggaatatt tgaaagtatc 1080
ccctcactcc aggaactagt gcttagctgg cagaactaat agccttcatt ggggcactag 1140
aattaggaga aggaaaaagg gtaaatatat atacagactc tgagtatgct cacctagtcn 1200
tccatgccca tgaggcaata tgcagagaaa gggaatteet aactteegag ggaacaecta 1260
tcacacatca ggaagccatt aggagattat tactggcagt acagaaacct aaagaggtgg 1320
aagtottaca otgotggggt catcagaaag gaaagaaaag ggaaatagaa gggaattgcc 1380
aagcagatat tgaagcaaaa agagctgcaa ggcaggaccc tc
                                                                  1422
<210> 23
<211> 2006
<212> DNA
<213> Homo sapiens
<400> 23
atgcagtggt cagtgataat ggaatacttg aaagtaatcc cctcactcca ggaactagtg 60
ctcagctagc agaactaata gccctcactt gggcactaga attaggagaa gaaaaaaggg 120
caaatatata tacagactot aaatatgott acctagtoot coatgoccat gcagcaatat 180
ggaaagaaag ggaattocta acttotgaga gaacacotat caaacatcag gaagcoatta 240
ggaaattatt attggctgta cagaaaccta aagaggtggc agtcttacac tgccggggtc 300
atcanaaagg aaaggaaagg gaaaatactt ttgcctgcaa ctatccaatg gaaattactt 360
aaaaccette atcaaacett teacttagge atcgatagea cecatcaaat ggccaaatca 420
ttatttactg gaccaggect tttcaaaact atcaagcaaa tattcagggc ctgtgaattq 480
tgccaaaaaa ataatcccct gcctcatcgc caagctcctt caggaaaaca aaaaacaggc 540
cattaccetg aaaaaaactg gcaactgatt tracceacaa gcccaaacct cagggatttc 600
agtatetaet agtetgggta aataetttea egggttggge aaaggeette eeetgtagga 660
cagaaaaggc ccaagaggta ataaaggcac tagttcatga aataattccc agattcggac 720
ttccccgagg cttacagagt gacaatagcc ctgctttcca ggccacagta acccagggag 780
tateccagge gttaggtata egatateaet tacaetgege etgaaggeca eagteeteag 840
ggaaggtcga gaaaatgaat gaaatactca aaggacatct aaaaaagcaa acccaggaaa 900
eccaceteae atggeetget etgttgeeta tageettaaa aagaatetge aaettteeee 960
aaaaagcagg acttagccca tacgaaatgc tgtatggaag gcccttcata accaatgacc 1020
tigigotiga occaagacag ocaacitagi igcagacato acciocitag ocaaatatoa 1080
```

```
acaagttott aaaacattac aaggaaccta toootgagaa gagggaaaag aactattoca 1140
cccttgtgac atggtattag tcaagtccct tctctctaat tccccatccc tagatacatc 1200
ctgggaagga ccctacccag tcattttatt taccccaact gcggttaaag tggctggagt 1260
ggtcttggat acatcacact tgagtcaaat cctggatact gccaaaggaa cctgaaaatc 1320
caggagacaa egetagetat teetgtgaac etetagagga titgegeetg etetteaaac 1380
aacaaccagg aggaaagtaa ctaaaatcat aaatccccca tggccctccc ttatcatatt 1440
tttctcttta ctgttctttt accetctttc actctcactg caceccetcc atgccgctgt 1500
atgaccagta geteceetta ecaagagttt etatggagaa tgeagegtee eggaaatatt 1560
gatgccccat cgtataggag tetttetaag ggaaccecca cetteactge ccacacceat 1620
atgecoegea actgetatea etetgecaet etitgeatge atgeaaatae teattattgg 1680
acaggaaaaa tgattaatee tagttgteet ggaggaettg gagteactgt etgttggaet 1740
tacttcaccc aaactggtat gtctgatggg ggtggagttc aagatcaggc aagagaaaaa 1800
catglaaaag aagtaatoto ccaactoaco egggtacatg gcacctctag coctacaaag 1860
gactagatet eteaaaacta catgaaacce teegtaeeca tactegeetg gtaageetat 1920
ttaataccac cctcactggg ctccatgagg tctcggccca aaaccctact aactgttgga 1980
tatgcctccc cctgaacttc aagcca
                                                                  2006
```

<210> 24 <211> 1948 <212> DNA

<213> Homo sapiens

<400> 24

actgcactct totggtccat gtttcttacg gctcgagctg agcttttgct caccgtccac 60 cactgetgtt tgccaccacc gcanacctge egetgactee catecetetg gateetgcag 120 ggtgteeget gtgeteetga teeagegagg egeceattge egeteecaat tgggetaaag 180 gcttgccatt gtncctgcac ggctaagtgc ctgggtttgt tctaattgag ctgaacacta 240 ntcactgggt tccatggttc tcttctgtga cccacggctt ctaatagaac tataacactt 300 accacatgge ccaagattee attesttgga atcegtgagg gcaagaacte caggtcagag 360 aatacgagge ttgccaccat cttggaageg geetgetace atettggaag tggttcacca 420 ccatcttggg agetetgtga gcaaggacce eceggtaaca ttttggcaac cacgaacgga 480 catccaaagt gatacatcct gggaaggacc ctacccagtc attttatcta ccccaactqc 540 ggttaaagtg gctggagtgg agtcttggat acatcacact tgagtcaaat cctggatact 600 gccaaaggaa cctgaaaatc caggagacaa cgctagctat tcctgtgaac ctctagagga 660 tttgcgcctg ctcttcaaac aacaaccagg aggaaagtaa ctaaaatcat aaatccccat 720 ggccctccct tatcatattt ttctctttac tgttgtttca ccctctttca ctctcactgc 780 accoccteca tgccgctgta tgaccagtag ctccccttac caagagtttc tatggagaat 840 gragegtere ggaaatattg atgececate gtataggagt ctttgtaagg gaaceeccac 900 cttcactgcc cacacccata tgccccgcaa ctgctatcac tctgccactc tttgcatgca 960 tgcaaatact cattattgga caggaaaaat gattaatcct agttgtcctg gaggacttgg 1020 agtcactgtc tgttggactt acttcaccca aactggtatg tctgatgggg gtggagttca 1080 agatcaggca agagaaaaac atgtaaaaga agtaatctcc caactcaccc gggtacatgg 1140 caccictage ecctacaaag gactagatet eteaaaaeta catgaaaece teegtaceca 1200 tactcgcctg gtaagcctat ttaataccac cctcactggg ctccatgagg tctcggccca 1260 aaaccctact aactgttgga tatgcctccc cctgaacttc aggccatatg tttcaatccc 1320 . tgtacctgaa caatggaaca acttcagcac agaaataaac accacttccg ttttagtagg 1380 acctcttgtt tccaatctgg aaataaccca tacctcaaac ctcacctgtg taaaatttag 1440 caatactaca tacacaacca actoccaatg catcaggtgg gtaactooto ccacacaaat 1500 agtotgocta cootcaggaa tattttttgt otgtggtaco toagoctato gttgtttgaa 1560 tggctcttca gaatctatgt gcttcctctc attcttagtg ccccctatgg ccatctacac 1620 tgaacaagat ttatacagtt atgtcatate taageeeege aacaaaagag tacceattet 1680 teettttgtt ataggageag gagtgetagg tgeactaggt actggeattg geggtateae 1740 aacctctact cagttctact acaaactatc tcaagaacta aatggggaca tggaacgggt 1800 egeegaetee etggteacet tgcaagatea aettaactee etageageag tagteettea 1860 aaatcgaaga gctttagact tgctaaccgc tgaaagaggg ggaacctgtt tatttttagg 1920 ggaagaatgc tgttattatg ttaatcaa

```
<210> 25
 <211> 1136
 <212> DNA
 <213> Homo sapiens
 <400> 25
 ccatggccat ctacactgaa caagatttat acagttatgt catatctaag ccccgcaaca 60
 aaagagtacc cattetteet titgttatag gageaggagt getaggtgea ctaggtactg 120
 geattggegg tateacaace tetacteagt tetactacaa actateteaa gaactaaatq 180
 gggacatgga acgggtcgcc gactccctgg tcaccttgca agatcaactt aactccctag 240
 cagcagtagt cetteaaaat egaagagett tagaeteget aacegetgaa agaggggaa 300
 cotgittatt titaggggaa gaatgotgit attatgitaa toaatoogga atogicactg 360
 agaaagttaa agaaattoga gatogaatao aacgtagago agaagagott ogaaacactg 420
 gaccetgggg cetectcage caatggatge cetggattet eccettetta ggacetetag 480
 cagetataat attgetacte etetttggae cetgtatett taaceteett gttaactttg 540
 tetettecag aategaaget gtaaaactae aaatggagee caagatgeag teeaagacta 600
 agatetaceg cagacecetg gaceggeetg etageceaeg atetgatgtt aatgacatea 660
 aaggcaccec teetgaggaa ateteagetg cacaacetet aetaegeeec aatteageag 720
gaagcagtta gagcggtcgt cggccaacct ccccaacagc acttaggttt tcctgttgag 780
 atgggggact gagagacagg actagetgga ttteetagge tgaetaagaa teectaagee 840
tagctgggaa ggtgaccaca tocaccttta aacacggggc ttgcaactta gttcacacct 900
gaccaatcag agagctcact aaaatgctaa ttaggcaaag acaggaggta aagaaatagc 960
caatcatcta ttgcatgaga gcacagcagg agggacaatg atcgggatat aaacccaagt 1020
cttcgagccg gcaacggcaa ccccctttgg gtcccctccc ttrgtatggg agctctgttt 1080
<210> 26
<211> 2782
<212> DNA
<213> Homo sapiens
<400> 26
atgggagetg titteatget atticaetet attaaatett geaactgeae tettetqqte 60
catgittett acggetegag etgagettit geteaccqte caccactqct gittigecace 120
accgcagace tgccgctgac teccatecet etggatectg cagggtgtee getgtgetee 180
tgatccageg aagegeecat tgeegeteec aattgggeta aaggettgee attgtteetg 240
cacggctaag tgcctgggtt tgttctaatt gagctgaaca ctagtcactg ggttccatgg 300
ttetettetg tgacccaegg ettetaatag aactataaca ettaccaeat ggeccaagat 360
tecatteett ggaateegtg aggecaacga actecaggte agagaataeg aagettgeea 420
ccatcttgga ageggeetge taccatettg gaagtggtte accaccatet tgggagetet 480
gtgagcaagg acccccggt gacattttgg cgaccaccaa cggacatccc aagtgataca 540
tcctgggaag gaccctaccc agtcatttta tctaccccaa ctgcggttaa agtggctgga 600
gtggagtett ggatacatea caettgagte aaateetgga taetgecaaa ggaacetgaa 660
aatccaggag acaacgctag ctattcctgt gaacctctag aggatttgcg cctgctcttc 720
aaacaacaac caggaggaaa gtaactaaaa tcataaatcc ccatgggcct cccttatcat 780
attititetet gtagigitet ticaccetgi ticactetea eigeaceece tecaigeege 840
tgtatgacca gtagctcccc tcacccagag tttctatgga gaatgcagcg tcccggaaat 900
attgatgece categratag gagtetttet aagggaacee ceaeetteae tgeceacace 960
catatgeece geaactgeta teactetgee actetttgea tgeatgeaaa tacteattat 1020
tggacaggaa aaatgattaa teetagttgt eetggaggae ttggagteae tgtetgttqq 1080
acttacttca cccaaactgg tatgtctgat gggggtggag ttcaagatca ggcaagagaa 1140
aaacatgtaa aagaagtaat ctcccaactc accggggtac atggcacctc tagcccctac 1200
aaaggactag atctctcaaa actacatgaa accctccgta cccatactcg cctggtaagc 1260
ctatttaata ccaccetcae tgggetecat gaggtetegg cccaaaacce tactaactgt 1320
tggatatgcc tccccctgaa cttcaggcca tatgtttcaa tccctgtacc tgaacaatgg 1380
aacaacttca gcacagaaat aaacaccact toogttttag taggacctct tgtttocaat 1440
```

```
gtggaaataa cccatacctc aaacctcacc tgtgtaaaat ttagcaatac tacatacaca 1500
 accaactece aatgeateag gtgggtaact ecteecacac aaatagtetg ectaecetea 1560
 ggaatatttt ttgtctgtgg tacctcagcc tatcgttgtt tgaatggctc ttcagaatct 1620
 atgtgcttcc totoattott agtgccccct atgaccatct acactgaaca agatttatac 1680
 agttatgtca tatctaagcc ccgcaacaaa agagtaccca ttcttccttt tgttatagga 1740
 gcaggagtgc taggtgcact aggtactggc attggcggta tcacaacctc tactcagttc 1800
 tactacaaac tatctcaaga actaaatggg gacatggaac gggtcgccga ctccctggtc 1860
 accttqcaaq atcaacttaa ctccctagca gcagtagtcc ttcgaaatcg aagagcttta 1920
 gacttgctaa ccgctgagag agggggaacc tgtttatttt taggggaaga atgctgttat 1980
 tatqttaatc aatccqqaat cqtcactqaq aaagttgaag aaattccaga tcgaatacaa 2040
 cqtataqcag aggagetteg aaacaetgga eeetggggee teetcageeg atggatgeee 2100
 tqqattctcc ccttcttagg acctctagca gctataatat tgctactcct ctttggaccc 2160
 tqtatctttg acctccttgt taactttgtc tcttccagaa tcgaagctgt gaaactacaa 2220
 atggagecca agatgeagte caagactaag atetacegca gacecetgga eeggeetget 2280
agcccacgat ctgatgttaa tgacatcaaa ggcacccctc ctgaggaaat ctcagctgca 2340
caacctctac tacgcccaa ttcagcagga agcagttaga gcggtggtcg gccaacctcc 2400
ccaacagcac traggttttc ctgttgagat gggggactga gagacaggac tagctggatt 2460
tcctaggctg actaagaatc cttaagccta ggtgggaagg tgaccacatc cacctttaaa 2520
caeggggett geaacttage teacacetga ceaateagag ageteactaa aatgetaatt 2580
aggcaaagac aggaggtaaa gaaatagcca atcatttatt geetgagage acagcaggag 2640
ggacaatgat cgggatataa acccaagttt tcgagccggc aacggcaacc ccctttgggt 2700
coordectt tgtatgggag ctctgttttc atgctatttc actctattaa atcttgcaac 2760
                                                                   2782
tgcaaaaaaa aaaaaaaaa aa
<210> 27
<211> 666
<212> DNA
<213> Homo sapiens
<400> 27
tgtccgctgt gctcctgatc cagcgaggcg cccattgccg ctcccaattg ggctaaaggc 60
ttgccattgt tcctgcacgg ctaagtgcct gggtttgttc taattgagct gaacactant 120
cactgggtte catggttete ttetgtgace caeggettet aatataacta taacacttac 180
cacatggccc aagattccat teettggaat eegtgaggec aagaacteca ggtcagagaa 240
tacgaggett gecaccatet tggaagegge etgetaceat ettggaagtg gttcaccace 300
atottgggag ototgtgago aaggacooco oggtaacatt ttggcaacca ogaacggaca 360
tocaaagtga atogaagotg taaaactaca aatggagooc aagatgcagt ccaagactaa 420
gatctacege agacccetgg aceggeetge tageccacga tetgatgtta atgacateaa 480
aggcacccct cctgaggaaa tctcagctgc acaacctcta ctacgcccca attcagcagg 540
aagcagttag ageggtegte ggecaacete eccaacagea ettaggtttt eetgttgaga 600
tgggggactg agagacagga ctagetggat tteetagget gaetaagaat eeetaageet 660
                                                                  666
agctgg
<210> 28
<211> 3372
<212> DNA
<213> Homo sapiens
<400> 28
gacttcccaa ataccagagg aagcagagtg gtttacagtc ctggaccttc aggatgcctt 60
cttctgcatc cctgtacatc ctgactctca attcttgttt gcctttgaag atacttcaaa 120
cocagcatet caacteacet ggactatttt accceaaggg tteagggata gteeceatet 180
atttggccag gcattagccc aagacttgag ccaatcctca tacctggaca cttgtccttc 240
ggtaggtgga tgatttactt ttggccgccc attcagaaac cttgtgccat caagccaccc 300
aagegetett caattteete getaeetgtg getaeatggt ttecaaacca aaggeteaac 360
totgotcaca goaggttact tagggotaaa attatocaaa ggoaccaggg cootcagtga 420
```

```
ggaacacate cagestatas tggettates teatescaaa accetaaage aactaaggg 480
 attectigge glaataggti tetgeegaaa atggatteee aggtatggeg aaatageeaq 540
 gtcattaaat acactaatta aggaaactca gaaagccaat acccatttag taagatggac 600
 aactgaagta gaagtggctt tocaggccct aacccaagcc ccagtgttaa gtttgccaac 660
 agggcaagac ttttgttcat atgtcacaga aaaaacagga atagctctag gagtccttac 720
 acagateega gggatgaget tgeaacetgt ggeacacetg actaaggaaa ttgatgtagt 780
ggcaaagggt tgacctcatt gtttacgggt agtggtggca qtaqcaqtct taqtatctqa 840
 agcagttasa ataatacagg gaagagatot tactgtqtqq acatotcatq atqtqaatqq 900
 catactcact gctaaaggag acttgtggct gtcagacaac tgtttactta aatgtcaggc 960
totattactt gaagggccag tgctgcgact gtgcacttgt gcaactctta acccaqccac 1020
 atttetteca gacaatgaag aaaagataaa acataactgt caacaagtaa tttetcaaac 1080
 ctatgccact cgaggggacc ttttagaggt tcctttgact gatcccgacc tcaacttqta 1140
tactgatgga agttcctttg tagaaaaagg acttcgaaaa gtggggtatg cagtggtcag 1200
tgataatgga atacttgaaa gtaatcccct cactccagga actagtgctc agctagcaga 1260
actaatagcc ctcacttggg cactagaatt aggagaagaa aaaaqggcaa atataataca 1320
gactetaaat atgettaeet agteeteeat geeeatgeag caatatggaa agaaagggaa 1380
ttcctaactt ctgagagaac acctatcaaa catcaggaag ccattaggaa attattattg 1440
getgtacaga aacctagaga ggtggcagtc ttacactgec ggggtcatca caaaggaaag 1500
gaaagggaaa tacaagagaa ctqccaaqca tatattqaaq ccaaaaqaqc tqcaaqqcaq 1560
gaccctccat tagaaatgct tattaaactt cccttagtat agggtaatcc cttccgggaa 1620
accaagecce agtactcage aggagaaaca qaatggggaa ceteaegagg cagttttete 1680
ccctcgggac ggttagccac tgaaqaaqqq aaaatacttt tqcctqcaac tatccaatqq 1740
aaattactta aaacccttca tcaaaccttt cacttaggca tcgatagcac ccatcagatg 1800
gccaaatcat tatttactgg accaggcctt ttcaaaacta tcaagcagat agtcagggcc 1860
tgtgaagtgt gccagagaaa taatcccctg ccttatcgcc aagctccttc aggagaacaa 1920
agaacaggee attaccetgg agaagactgg caactgattt tacccacaag cccaaacete 1980
agggatttca gtatctacta gtctgggtag atactttcac gggttgggca gaggccttcc 2040
cctgtaggac agaaaaggcc caagaggtaa taaaggcact agttcatgaa ataattccca 2100
gatteggact teecegagge ttacagagtq acaatageee tgettteeag gecacagtaa 2160
cocagggagt atcccaggeg traggtatac gatatcactt acactgegee tgaaggeeac 2220
agtecteagg gaaggtegag aaaatgaatq aaacacteaa aggacateta aaaaagcaaa 2280
cccaggaaac ccacctcaca tggcctqttc tgttqcctat agccttaaaa agaatctgca 2340
actttcccca aaaagcagga cttaqcccat acqaaatqct gtatqqaagg cccttcataa 2400
ccaatgacet tgtgcttgac ccaagacage caacttagtt gcagacatca cetecttage 2460
caaatatcaa caagttotta aaacattaca aggaacctat cootgagaag aggaaaagaa 2520
tattccaccc aagtgacatg gtattagtca agtcccttcc ctctaattcc ccatccctag 2580
atacatectg ggaaggacee tacccagtea ttttatetac cecaactgeg gttaaagtgg 2640
ctggagtgga gtcttggata catcacactt gagtcaaatc ctggatactg ccaaaggaac 2700
ctgaaaatcc aggagacaac gctagctatt cctgtgaacc tctagaggat ttgcgcctgc 2760
tottcaaaca acaaccagga ggaaaaatcq aagctgtaaa actacaaatg gagcccaaga 2820
tgcagtccaa gactaagatc taccgcagac ccctggaccg gcctgttagc ccacgatctg 2880
atgitaatga catcaaaggc acceptectg aggaaatete agetgeacaa cetetactac 2940
geoceaatte ageaggaage agttagageg gtegteggee aaceteeeca acageaetta 3000
ggttttcctg ttgagatggg ggactgagag acaggactag ctggatttcc taggctgatt 3060
aagaateeet aageetaget gggaaggtga ceacateeae etttaaacae ggggettgea 3120
acttagetea cacetgacea atcagagage teactaaaaat getaattagg caaagacagg 3180
aggtaaagaa atagccaatc atttattgcc tgagagcaca gcaggaggga caatgatcgg 3240
gatataaacc caagttttcg agccggcaac ggcaaccccc tttgggtccc ctccctttgt 3300
atgggagete tgtttteatg etattteact etattaaate ttgeaactge aaaaaaaaaa 3360
aaaaaaaaa aa
```

<210> 29

<211> 2372

<212> DNA

<213> Homo sapiens

<400> 29

```
actgcactet tetggtccat gtttettacg getegagetg agettttget cacegtecae 60
  cactgotgtt tgccaccacc gcagacctgc cgctgactcc catccctctg gatcctgcaq 120
 ggtgtccgct gtgctcctga tccagcgagg cgcccattgc cgctcccaat tgggctaaag 180
 gcttgccatt gttcctgcac ggctaagtgc ctgggtttgt tctaattgag ctgaacacta 240
 atcactgggt tocatggttc tottotgtga occacggott ctaatagaac tataacactt 300
 accacatgge ccaagattee attecttgga atccgtgagg ccaagaacte caggteagag 360
 aatacgaggc ttgccaccat cttggaagcg gcctgctacc gtcttggaag tggttcacca 420
 ccatcttggg agctctgtga gcaaggaccc cccggtaaca ttttggcaac caacgacgga 480
 catccaaagt gatgggaaac gttccccgca agacaaaaac gcccctaaga cgtattctgg 540
 agaattggga ccaatttgac cctcagacac taagaaagaa acgacttata ttcttctgca 600
 gtgccgcctg gcactcctga gggaagtata aattataaca ccatcttaca gctagacctc 660
 ttttgtagaa aaggcaaatg gagtgaagtg ccataagtac aaactttctt ttcattaaga 720
 gacaactcac aattatgtaa aaagtgtgat ttatgcccta caggaagcct tcagagtcta 780
 ectecetate ccageatece egacteette eccaactaat aaggaceece etteaaceca 840
 aatggtccaa aaggagatag acaaaagggt aaacagtgaa ccaaagagtg ccaatattcc 900
 ccaattatga cccctccaag cagtgggagg aagagaattc ggcccagcca gagtgcatgt 960
 gcctttttct ctcccagact taaagcaaat aaaaacagac ttaggtaaat tctcagataa 1020
 CCCtgatggc tatattgatg ttttacaagg gttaggacaa ttctttgatc tgacatqgaq 1080
 agatataatg teactgetaa atcagacaet aaceecaaat gagagaagtg ceaccataae 1140
 tgcagcctga gggtttggcg tctctggtat ctcagtcagg tcaatggata nggatgacaa 1200
 cagaaggaaa ganaatgatt ccccacagge cagcaggcag ttcccagtct agaccctcat 1260
 tgggacacag aatcagaaca tggagattgg tgctgcagac atttgctaac ttgtgtgcta 1320
 gaaggactaa ggaaaactag gaagaagtet atgaattact caatgatgte caccataaca 1380
 cagggaaggg aagaaaatcc tactgccttt ctggagagac taagggaggc attgaggaag 1440
 cgtgcctctc tgtcacctga ctcttctgaa ggccaactaa tcttaaagcg taagtttatc 1500
 actcagtcag ctgcagacat tagaaaaaac ttcaaaaagtc tgccgtaggc ccggagcaaa 1560
acttagaaac cctattgaac ttggcaacct cggtttttta taatagagat caggaggagc 1620
 aggeggaaca ggacaaacgg gattaaaaaa aaggecaceg etttagteat gacceteagg 1680
caagtggact ttggaggctc tggaaaaggg aaaagctggg caaattgaat gcctaatagg 1740
gettgettee agtgeggtet acaaggacac tttaaaaaaag attgteeaag tagaagtaag 1800
degecentic greeatgees ettatticaa gggaateast ggaaggeesa etgececagg 1860
ggacaaaggt cttttgagtc agaagccact aaccagatga tccagcagca ggactgaggg 1920
tgcctggggc aagcgccatc ccatgccatc accctcacag agccctgggt atgcttgacc 1980
attgagggee aggaaggttg teteetggae actggtgegg tettettagt ettactette 2040
tgtcccggac aactgtcctc cagatctgtc actattctga gggggtccnt aagacgggca 2100
gtcactagat acttttccc agccactaag ttatgaactg gggagcttta ttctttcac 2160
atgettttet aattatgett gaaageeeca etacettgtt agggagagae attetageaa 2220
aagcaggggc cattatacac ctgaacataq qagaaqqaac acccqtttqt tgtncccctg 2280
cttgaggaag gaattaatcc tgaagtctgg gcaacagaag gacaatatgg acgagccaaa 2340
gaatgcccgt cctgttcaag ttaaactaaa gg
<210> 30
<211> 7582
<212> DNA
<213> Homo sapiens
<400> 30
caacaategg gatataaace caggeatteg agetggeaac ageageeeer etttgggtee 60
cttccctttg tatgggagct gttttcatgc tatttcactc tattaaatct tgcaactgca 120
ctcttctggt ccatgtttct tacggctcga gctgagcttt tgctcaccgt ccaccactgc 180
tgtttgccac caccgcanac etgecgetga etcecatece tetggatect geagggtgte 240
egetgtgete etgateeage gargegeeea ttgeegetee caattggget aaaggettge 300
cattgtncct gcacggctaa gtgcctgggt ttgttctaat tgagctgaac actantcact 360
gggttccatg gttctcttct gtgacccacg gcttctaata kaactataac acttaccaca 420
tggcccaaga ttccattcct tggaatccgt gaggscaacg aactccaggt cagagaatac 480
gargettgee accatettgg aageggeetg etacertett ggaagtggtt caccaccate 540
```

ttgggagete tgtgageaag gaceeeegg tracattttg geraeeamsr aeggaeatee 600

maagtgatgg gaaacgttcc ccgcaagaca aaaacgcccc taagacgtat tctggaraat 660 tgggamcaat ttgaccetca gacactaaga aagaaacgae ttatattett etgeagtgee 720 gcctggcact cctgagggaa gtataaatta taacaccatc ttacagctag acytcttttg 780 tagaaaaggc aaatggagtg aagtgccata agtacaaact ttcttttcat taagagacaa 840 ctcacaatta tgtaaaaagt gtgatttatg coctacagga agccttcaga gtctacctcc 900 ctateccage ateccegaet cettecceam ytaataagga eccecettea acceaaatgg 960 tecaaaagga gatagacaaa agggtaaaca gtgaaccaaa gagtgecaat attececaat 1020 tatgacccct cccaagcagt gggaggaaga gaattcggcc cagccagagt gcatgtgcyt 1080 tttyytotoo cagactraaa gcaaataaaa acagactrag gtaaattoto agataayoot 1140 gatggctata ttgrtgtttt acaagggtta ggacaattct ttgatctgac atggagagat 1200 atatatgtca ctgctaaatc agacactaac cccaaatgag agaagtgcca ccataactgc 1260 agectgagrg tttggegate tetggtatet cagteaggte aatggatang gatgacaaca 1320 gaaggaaaga naatgattoc ccacaggcca gcargcagtt cccagtctas accetcattq 1380 gggacacaga aatcagtaac atgggagatt ggtgctgcag acatttgcta acttgtgtgc 1440 tasaaggact aaggaaaact asgaagaaar totaygaatt actcaatgat gtccaccata 1500 acacagggga agggaagaaa atcctactgc ctttctggag agactaaggg aggcattgag 1560 gaagegtgee tetetgteae etgaetette tgaaggeeaa etaatettaa agegtaagtt 1620 tatcactcag tcagctgcag acattagaaa aaacttcaaa agtctgccgt aggcccggag 1680 caaaacttag aaaccctatt gaacttggca acytcggttt tttataatag agatcaggag 1740 gagcaggcgg aacaggacaa acgggattaa aaaaaaggcc accgctttag tcatgaccct 1800 caggcaagtg gactttggag gctctggaaa agggaaaagc tgggcaaatt gaatgcctaa 1860 tagggettge ttecagtgeg gtetacaagg acaetttaaa aaagattgte caagtagaag 1920 taagccgccc cttcgtccat gccccttatt tcaagggaat cactggaagg cccactgccc 1980 caggggacaa aggtcttttg agtcagaagc cactaaccag atgatccagc agcaggactg 2040 agggtgcctg gggcaagegc cateceatge cateaceete acagagecet gggtatgett 2100 gaccattgag ggccaggaag gttgtctcct ggacactggt gcggtcttct tagtcttact 2160 cttctgtccc ggacaactgt cctccagatc tgtcactatt ctgagggggt ccntaagacg 2220 ggcagtcact agatacttty teccagecac taagttatga actggggage tttattettt 2280 tcacatgctt ttctaattat gcttgaaagc cccactacct tgttagggag agacattcta 2340 gcaaaagcag gggccattat acacctgaac ataggagaag gaacacccgt ttgttgtncc 2400 cctgcttgag gaaggaatta atcctgaagt ctgggcaaca gaaggacaat atggacgagc 2460 caaagaatgc ccgtcctgtt caagttaaac taaaggattc cacttccttt ccctaccaaa 2520 ggcagtaccc cctcagaccc aaggcccaac aaggattcca aaagattgtt aaggacttaa 2580 aagcccaagg cttagtaaaa ccatgcataa ctccctgcag taattccgta gtggattgag 2640 gaggcacaga aacccagtgg acagtggagg gttagtgcaa gatctcagga ttatcaatgg 2700 aggoogttgt cottttatac coagotgtac ctagocotta tactqtqmyt toccaaatac 2760 cagaggaagc agagtggttt acastcctgg accttmagga tgccttcttc tgcatccctg 2820 tacatectga eteteaatte ttgtttgeet ttgaagatae tteaaaceca reateteaae 2880 tracetggad tritttadec caagggitea gggatagyde ceatetaitt ggeeaggeat 2940 tagcccaaga cttgagycar tymtcatacc tggacactct tgtccttcrg takgtggatg 3000 atttactttt rgcygccyrt tcagaaacct tgtgccatca agccacccaa gcrctcttma 3060 attteetege yacctgtgge tacawggttt ccaaacsara rgctearete tgeteacage 3120 aggttaaata cttaggrcta arattatcca aaggcaccar ggccctcagt gaggaayrya 3180 tocagoctat actggottat cotcatoyca aaaccotaaa gcaactaagr grrttoottg 3240 gcrtaayagg yttctgccga awatggattc cccaggtwtg gcraaatagc caggycatta 3300 watacastaa ttaaggaaac tcagaaagcc aatacccatt tartaagatg gayamctgaa 3360 gymraagtgg ctttccaggc ccctaaagaa ggccttaaac ccaagyccca gtgttaagyt 3420 tgccaacrgg gcaagacttt tsttyatayr tcacagaaaa aaacagraay agctctrgga 3480 gtccttacac agrtccragg gaygagettg caaccyrtgg cryacetgas taaggaaavt 3540 gatgtagtgg caaagggttg rcytcattgt ttaygggtag tggtggcagt agcagtykta 3600 gtatctgaag cagttaaaat aatacagggr agagatctta ctgtgtggac atctcatgak 3660 gtgaayrgca tactcactgc taaaggagac ttgtggctgt cagacaacyg tttacttaaa 3720 trtcaggete tattacttga arggecagtg etgeractgt geacttgtge aactettaac 3780 ccagycncat ttcttccaga caatgaagaa aagataraay ataactgtca acaartaatt 3840 totcaaacct atgocactog aggggacctt ytagargtto cyttgactga tocygacctt 3900 caacttgtat actgatggaa gttcctttgt agaaaaagga cttcgaaaag yggggtatgc 3960 agtggtcagt gataatggaa tayttgaaag taatcccctc actccaggaa ctagtgctya 4020 gctrgcagaa ctaatagccy tcaytkgggc actagaatta ggagaagraa aaagggyaaa 4080

tatatataca gactetrart atgetyacet agtentecat geceatgmrg caatatgsar 4140 agaaagggaa ttcctaactt cygagrgaac acctatcama catcaggaag ccattaggar 4200 attattaytg gcwgtacaga aacctaraga ggtggmagtc ttacactgcy ggggtcatca 4260 naaaggaaag raaagggaaa tasaagrgaa ytgccaagca katattgaag cmaaaagagc 4320 tgcaaggcag gacctccat tagaaatgct tattaaactt cccttagtat agggtaatcc 4380 cttccgggaa accaagcccc agtactcagc aggagaaaca gaatggggaa cctcacgagg 4440 cagtititite coctogggac ggitagocac tgaagaaggg aaaatactit tgcctgcaac 4500 tatocaatgg aaattactta aaaccottca tcaaaccttt cacttaggca tcgatagcac 4560 ccatcaratg gccaaatcat tatttactgg accaggcctt ttcaaaacta tcaagcarat 4620 akteaqqqce tqtqaaktqt gecararaaa taateecetg cetyategee aageteette 4680 aggaraacaa araacaggcc attaccctgr araaractgg caactgattt tacccacaag 4740 cccaaacete agggatttea gtatetaeta gtetgggtar ataettteae gggttgggea 4800 raggeettee cetgtaggae agaaaaggee caagaggtaa taaaggeact agtteatgaa 4860 ataattccca gattcggact tccccgaggc ttacagagtg acaatagccc tgctttccag 4920 gccacagtaa cccagggagt atcccaggcg ttaggtatac gatateactt acactgcgcc 4980 tgaaggccac agtcctcagg gaaggtcgag aaaatgaatg aaayactcaa aggacatcta 5040 aaaaagcaaa cccaggaaac ccacctcaca tggcctgytc tgttgcctat agccttaaaa 5100 agaatetgea aettteecea aaaageagga ettageeeat aegaaatget gtatggaagg 5160 cccttcataa ccaatgacct tgtgcttgac ccaagacagc caacttagtt gcagacatca 5220 cctccttagc caaatatcaa caagttctta aaacattaca aggaacctat ccctgagaag 5280 agggaaaaga actattccac ccwwgtgaca tggtattagt caagtccctt cyctctaatt 5340 ccccatcct agatacatcc tgggaaggac cctacccagt cattttatyt accccaactg 5400 cggttaaagt ggctggagtg gagtcttgga tacatcacac ttgagtcaaa tcctggatac 5460 tgccaaagga acctgaaaat ccaggagaca acgctagcta ttcctgtgaa cctctagagg 5520 atttgcgcct gctcttcaaa caacaaccag gaggaaagta actaaaatca taaatccccc 5580 atggsected ottateatat tittetetkt astgitstit yaccetstit cacteteact 5640 gcacccctc catgccgctg tatgaccagt agctccccty accmagagtt tctatggaga 5700 atgcagcgtc ccggaaatat tgatgcccca tcgtatagga gtctttstaa gggaaccccc 5760 accttcactg cccacacca tatgccccgc aactgctatc actctgccac tetttgcatg 5820 catqcaaata ctcattattq qacaqqaaaa atgattaatc ctagttgtcc tggaggactt 5880 ggagtcactg tctgttggac ttacttcacc caaactggta tgtctgatgg gggtggagtt 5940 caagatcagg caagagaaaa acatgtaaaa gaagtaatct cccaactcac csgggtacat 6000 ggeaceteta geceetacaa aggaetagat eteteaaaae tacatgaaae eeteegtace 6060 catactegee tggtaageet atttaatace acceteactg ggetecatga ggteteggee 6120 caaaacccta ctaactgttg gatatgcctc cccctgaact tcargccata tgtttcaatc 6180 cctgtacctg aacaatggaa caacttcagc acagaaataa acaccacttc cgttttagta 6240 ggacctcttg tttccaatst ggaaataacc catacctcaa acctcacctg tgtaaaattt 6300 agcaatacta catacacaac caacteecaa tgeateaggt gggtaactee teecacacaa 6360 atagtctgcc taccctcagg aatatttttt gtctgtggta cctcagccta tcgttgtttg 6420 aatggctctt cagaatctat gtgcttcctc tcattcttag tgcccccyat grccatctac 6480 actgaacaag atttatacag ttatgtcata tctaagcccc gcaacaaaag agtacccatt 6540 cttccttttg ttataggagc aggagtgcta ggtgcactag gtactggcat tggcggtatc 6600 acaaceteta etcagtteta etacaaaeta tetcaagaae taaatgggga catggaaegg 6660 gtegecgaet ceetggteac ettgeaagat caaettaaet eeetageage agtagteett 6720 craaatcgaa gagctttaga ctygctaacc gctgaragag ggggaacctg tttattttta 6780 ggggaagaat gctgttatta tgttaatcaa tccggaatcg tcactgagaa agttraagaa 6840 attcsagatc gaatacaacg takagcagar gagcttcgaa acactggacc ctggggcctc 6900 ctcagccrat ggatgccctg gattctcccc ttcttaggac ctctagcagc tataatattg 6960 ctactcctct ttgqaccctg tatetttrac ctecttgtta actttgtctc ttccageatc 7020 gaagetgtra aactacaaat ggageecaag atgeagteea agaetaagat etaeegeaga 7080 cccctggacc ggcctgytag cccacgatct gatgttaatg acatcaaagg cacccctcct 7140 gaggaaatot cagotgoaca acototacta ogococaatt cagoaggaag cagttagago 7200 ggtsgtcggc caacctcccc aacagcactt aggttttcct gttgagatgg gggactgaga 7260 gacaggacta gctggatttc ctaggctgay taagaatccy taagcctags tgggaaggtg 7320 accacateca cetttaaaca eggggettge aacttagyte acacetgace aatcagagag 7380 ctcactaaaa tgctaattag gcaaagacag gaggtaaaga aatagccaat catytattgc 7440 mtgagagcac agcaggaggg acaatgatcg ggatataaac ccaagtyttc gagccggcaa 7500 eggeaacece etttgggtee eetecetttg tatgggaget etgtttteat getattteae 7560

totattaaat ottgoarctg or

7582

<210> 31

<211> 363

<212> PRT

<213> Homo sapiens

<400> 31

Met Gly Asn Ile Pro Pro Lys Ala Lys Thr Pro Leu Arg Cys Ile Leu

Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu Arg Lys Lys Arg Phe

lle Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr Pro Leu Gln Gly Arg

Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr Asn Ile Ile Leu Gln 55

Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp Ser Glu Val Pro Tyr

Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser Gln Leu Cys Lys Lys

Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser Pro Pro Pro Tyr Pro 105

Ser Val Pro Pro Pro Thr Pro Ser Ser Thr Asn Lys Asp Pro Pro Leu 120

Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly Val Asn Asn Glu Pro

Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro Leu Gln Ala Val Arg Gly 155

Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro Phe Ser Leu Ser Asp 170

Leu Lys Gln Ile Lys Ile Asp Leu Gly Lys Phe Ser Asp Asn Pro Asp 185

Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln Ser Phe Asp Leu Thr 195

Trp Arg Asp Ile Met Leu Leu Leu Asn Gln Thr Leu Thr Pro Asn Glu 215

Arg Ser Ala Ala Val Thr Ala Ala Arg Glu Phe Gly Asp Leu Trp Tyr 230

Leu Ser Gln Val Asn Asn Arg Met Thr Thr Glu Glu Arg Thr Thr Pro 250

Thr Gly Gln Gln Ala Val Pro Ser Val Asp Pro His Trp Asp Thr Glu

Ser Glu His Gly Asp Trp Cys His Lys His Leu Leu Thr Cys Val Leu 280

Glu Gly Leu Arg Lys Thr Arg Lys Lys Pro Met Asn Tyr Ser Met Met 295

Ser Thr Ile Thr Gln Gly Lys Glu Glu Asn Pro Thr Ala Phe Leu Asp 310

Arg Leu Arg Glu Ala Leu Arg Lys His Thr Ser Leu Ser Pro Asp Ser

Ile Glu Gly Gln Leu Ile Leu Lys Asp Lys Phe Ile Thr Gln Ser Ala

Ala Asp Ile Arg Lys Asn Phe Lys Ser Leu Pro 360